EPIGENETIC REGULATION OF *Hoxa1* AND *Hoxa2*

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

in the Pharmacy Graduate Program

University of Saskatchewan

Saskatoon, Saskatchewan

Canada

Ran Bi

©Copyright Ran Bi, January 2017, All rights reserved.

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is also understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis. Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Chair of the Pharmacy and Nutrition Graduate Program
College of Pharmacy and Nutrition
University of Saskatchewan
104 Clinic Place
Saskatoon, SK, Canada, S7N 2Z4

ABSTRACT

Hoxa1 and *Hoxa2* are master regulators in the development of hindbrain, ear, palate, bone and cardiovascular development. There is little information on the epigenetic regulator(s) of *Hoxa2* gene during development. In this thesis, I have determined whether regulation of *Hoxa2* is occurring via a specific epigenetic pathway, and investigated the role of DNA methylation, noncoding RNAs (microRNAs and long non-coding RNAs) and histone protein modification.

First, analysis of *Hoxa2* promoter revealed the presence of three CpG islands near the *Hoxa2* 5′ regulatory region. Using methylation specific PCR (MSP) and the bisulfite specific PCR (BSP) primers followed by DNA sequencing, I found the methylation status of CpG island 1 remains unmethylated and that the DNA methylation status of the *Hoxa2* promoter does not change with the spatio-temporal expression of *Hoxa2* during palatogenesis. These findings indicate that DNA methylation does not appear to play a key role in the epigenetic regulation of *Hoxa2* gene during palatogenesis.

My second objective was to determine whether specific miRNAs impact *Hoxa2* expression. I performed in-silico analysis and identified six miRNAs that have the potential to bind 3'UTR of the *Hoxa2* gene. The miR-669b and miR-376c were capable of down regulating *Hoxa2* expression at both transcriptional and translational level. Two direct miR-669b binding sites were identified on mouse *Hoxa2* 3'UTR. Luciferase assays showed that the two miR-669b binding sites appear to work independently of each other and that mutations within the seed sequences abrogated luciferase activity. I further analyzed the degree of sequence similarity of both miR-669b binding sites and found that binding site 1 is evolutionarily conserved between the five species (human, mouse, rat, chimpanzee and dog). In the developing mouse palate (from E13 to E15), miR-669b showed a complementary expression to that of *Hoxa2*. No direct interaction between miR-376c

and *Hoxa2* 3'UTR was identified. Thus my results indicated that the miR-669b likely plays a role in regulating *Hoxa2* expression during palate development

My third objective was to characterize a new lncRNA (*mHotairm1*) that I identified between mouse *Hoxa1* and *Hoxa2* intergenic region. I demonstrated that *mHotairm1* is involved in recruiting MLL1/WDR5 to *Hoxa1* and *Hoxa2* genes and regulating their expressions. *In situ* hybridization histochemistry of E14 developing palate showed expression of *mHotairm1* in medial edge epithelia (MEE), indicating *mHotairm1* may play a role in the palatal fusion. Downregulation of *mHotairm1* in NIH 3T3 cells resulted in significantly decreased expression of *Hoxa1* and *Hoxa2* expression, whereas treatment with ATRA resulted in increased expression of *mHotairm1*, *Hoxa1* and *Hoxa2*. Using capture hybridization analysis of RNA targets (CHART) and pull down assays, I found that the TrxG protein WDR5 is associated with *mHotairm1*, and knockdown of *mHotairm1* resulted in reduced occupancy of gene activation mark H3K4me3 and increased occupancy of gene suppression mark H3K27me3, suggesting MLL1/WDR5 complex may be playing a role in the regulation of *Hoxa1* and *Hoxa2* gene expression through *mHotairm1*.

Lastly, I found that WDR5 was sumoylated. This modification appears to be important for its interaction with *mHotairm1* and MLL and for its cellular distribution, primarily to the nuclei. Following ATRA treatment, although the total WDR5 protein remained unchanged, an increase in sumoylated WDR5 was observed together with increased expression of *mHotairm1*, *Hoxa1* and *Hoxa2* gene. These findings reveal that sumoylated WDR5 with its interaction with *mHotairm1* plays an important role in H3K4me3 and H3K27me3 occupancy and influencing the epigenetic regulation of *Hoxa1* and *Hoxa2* genes.

ACKNOWLEDGEMENTS

First I would like to acknowledge the support of my supervisor Adil J. Nazarali. Without him I would never have such great chance to study abroad. During my study at University of Saskatchewan, he gave me great support and guidance in both my academic and personal life.

I would also like to thank my external examiner Dr. Yuewen Gong for attending my defence and my committee members, Dr. Jane Alcorn, Dr. Ildiko Badea, Dr. Stanley Moore and Dr. Yuliang Wu, for their suggestions throughout my PhD study. This thesis would not have been possible without their support.

I would like to give my thanks to my lab members, especially Shaoping Ji. He generously provided me suggestions and supports on my experiments.

DEDICATION

This thesis is dedicated to all my friends and family. In particular, I would like to dedicate this thesis to my parents, Chaoyuan Bi and Liying Wei who supported me mentally and financially during my study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3' UTR – 3' Untranslated Region

5-Aza-CdR – 5-aza-20-Deoxycytidine

5hmC – 5-Hydroxymethylcytosine

5mC – 5-Methylcytosine

A – Adenine

Ab/Am – Antibiotic/Antimycotic

Abd-A – Abdominal-A

Abd-B – Abdominal-B

ABDS – Athabascan Brainstem Dysgenesis Syndrome

AdoMet or SAM – S-Adenosyl-l-Methionine

AGO – Argonaute

ANT-C – Antennapedia Complex

AP – Anterior-Posterior

ASD – Autism Spectrum Disorder

ASH1 – Absent Small or Homeotic Discs 1

ASH2L – Absent Small or Homeotic 2-Like

ATRA – All-*trans* Retinoic Acid

BCS – Bovine Calf Serum

BD – Bromo Domain

BER – Base Excision and Repair

BIG-3 – BMP-2 Induced Gene 3kb

Bmp7 – Bone Morphogenetic Protein 7

BSA – Bovine Serum Albumin

BSAS – Bosley–Salih–Alorainy Syndrome

BSP – Bisulfite Specific PCR

BX-C – Bithorax Complex

C – Cytosine

CBP – CREB-Binding Protein

CHART – Capture hybridization analysis of RNA targets

CHD8 – Chromodomain, Helicase, DNA-binding 8

ChIP – Chromatin Immunoprecipitation

CNCC – Cranial Neural Crest Cell

CNS – Central Nervous System

CREB – cAMP-Response Element Binding

CxxC – Cysteine x x Cysteine (x – any amino acid)

D – Aspartic Acid

DIG – Digoxigenin

DNMT – DNA Methyltransferase

DMEM – Dulbecco's Modified Eagle Medium

E – Embryonic Day

E1A – Early-Region 1A

EOC – Epithelial Ovarian Cancer

ES – Embryonic Stem

ESC – Embryonic Stem Cell

EZH2 – Enhancer of Zeste Homolog 2

FAM – Fluorescein Amidite

FGF1 – Fibroblast Growth Factor 1

FGF18 – Fibroblast Growth Factor 18

Fgfr3 – Fibroblast Growth Factor Receptor 3

Foxd3 – Forkhead Box D3

gDNA – Genomic DNA

GNAT – GCN5-related N-acetyltransferase

H3K4me3 – Trimethylation of Histone 3 lysine 4

H3K4me2 – Dimethylation of Histone 3 lysine 4

H3K27me3 – Trimethylation of Histone 3 lysine 27

H3T11P – Histone 3 Threonine 11 Phosphorylation

HAT– Histone Acetyltransferase

HDAC – Histone Deacetylase

HKMT – Histone Lysine Methyltransferase

hLuc – Firefly Luciferase

HMTase – Histone Methyltransferase

Hmx1 – H6 Family Homeobox 1

Hnf1b – Hepatocyte Nuclear Factor 1-Beta

HOM-C – Homeotic Complex

hRLuc – Renilla Luciferase

HRP – Horse Radish Peroxidase

HUVEC – Human Umbilical Vein Endothelial Cell

IFN – Type I Interferons

IGF1R – Insulin-like growth factor1 receptor

IP – Immunoprecipitation

ISH – In Situ Hybridization

ISWI – Imitation Switch

iPS – Induced Pluripotent Stem

JmjC – Jumonji C

KAT– Lysine Acetyltransferase

KDAC – Lysine Deacetylase

KDM – Lysine Demethylase

lacZ – β-galactosidase

Lhx5 – LIM Homeobox 5

LncRNA – Long Noncoding RNA

MBD – Methyl-CpG Binding Domain Protein

MBP – Methyl-CpG Binding Protein

MEE – Medial Edge Epithelia

MEPM – Mouse Embryonic Palate Mesenchyme

MES – Medial Edge Seam

mHotairm1 – Mouse Hotairm1

MLL1 – Mixed Lineage Leukemia 1

MOF – Male Absent On the First

miRISC – miRNA-Containing RNA-Induced Silencing Complex

miRNA – microRNA

mRNA – Message RNA

MSP – Methylation Specific PCR

Msx1 – Msh Homeobox 1

Nab1 – NGFI-A Binding Protein 1

NEM – N-Ethylmaleimide

NCOA – Nuclear Receptor Co-activator

NPC – Nuclear Pore Complex

NSL – Nonspecific Lethal

nt – Nucleotide

ORF – Open Reading Frame

PA – Pharyngeal Arch

Pax9 – Paired Box 9

PBA – 4-Phenylbutyric Acid

PcG – Polycomb Group

PHB – PHABULOSA

PHD – Plant Homeodomain

PRC – Polycomb Repressive Complex

PRE – PcG Response Element

pre-miRNA – miRNA Precursors

pri-miRNA – Primary transcript of miRNA

PRMT6 – Protein Arginine Methyltransferase 6

pRNA – Promoter-Associated RNA

PTC – Papillary Thyroid Carcinoma

PTH – Parathyroid Hormone

qRT-PCR – Quantitative Real-Time PCR

RACE – Rapid Amplification of cDNA Ends

RBBP5 – Retinoblastoma Binding Protein 5

rDNA – Ribosomal DNA

RIPA – Buffer Radioimmunoprecipitation Assay Buffer

RT – Reverse Transcription

RQ – Relative Quantity

SAE – SUMO Activation Enzyme

SAR – Ring Finger-Associated

SET – Suppressor of Variegation-Enhancer of Zeste-Trithorax

snRNA – Small Nuclear RNA

Sox6 – SRY (Sex Determining Region Y)-Box 6

SRC – Steroid Receptor Coactivator

SSC – Saline Sodium Citrate

SUMO – Small Ubiquitin-Like Modifier

SWI/SNF – Switch/Sucrose Nonfermentable

T – Thymine

TAD – Transactivation Domain

TET – Ten-eleven Translocation

TDG – Thymine DNA Glycosylase

TF – Transcription Factors

TGFβ – Transforming Growth Factor β

TRE – TrxG Response Element

TrxG – Trithorax Group

Tsg – Twisted Gastrulation

U – Uracil

Ubx – Ultrabithorax

Uchl1 – Ubiquitin Carboxyterminal Hydrolase L1

VISA – Virus-Induced Signaling Adaptor

VNC – Ventral Nerve Cord

W – Tryptophan

WDR5 – WD Repeat Domain 5

WDS – WDR5 in *Drosophila*

Wnt5a – Wnt Family Member 5a

XCI – X-Chromosome Inactivation

Xi – Inactive X-chromosome

Xa – Active X-chromosome

Zic1 – Zic Family Member 1

CHAPTER 1

1. Summary

Hox genes are evolutionarily conserved homeodomain-containing transcription factors that specify cell identity in early development (McGinnis and Krumlauf, 1992; Banerjee-Basu and Baxevanis, 2001; Mallo and Alonso, 2013). In addition to their role in embryo development, *Hox* genes are also involved in different types of human diseases and cancer including leukemia (Alharbi *et al*., 2013) and breast cancer (Bhatlekar *et al*., 2014). In vertebrates, 39 *Hox* genes are organized into four Hox gene clusters (HoxA-D) and they encode Hox proteins that regulate specific morphological diversity along the anterior-posterior (AP) axis (Akin and Nazarali, 2005; De Kumar and Krumlauf, 2016.) *Hoxa1* and *Hoxa2* are the first *Hox* genes to be expressed during embryonic development (Murphy and Hill, 1991). *Hoxa1* plays a role in hindbrain, inner ear and cardiovascular development (Makki and Capecchi, 2011; Qiao *et al.,* 2015; Makki and Capecchi, 2012) and *Hoxa2* is important for the development of inner ear, palate, hindbrain and bone (Minoux *et al*., 2013; Smith *et al*., 2009; Gavalas *et al*., 1997; Kitazawa *et al*., 2015). The proper expression of *Hox* genes is important in many biological procesesses, making regulation of *Hox* genes critical to living organisms. There is little information available on the mechanism(s) involved in the regulation of *Hoxa2* gene expression. Hence, my research has involved in an investigation of the epigenetic regulation *Hoxa2* gene expression during mouse palate development and in the NIH 3T3 cell line.

Epigenetics refers to the heritable changes in gene expression without any changes in the DNA sequence. Regulation of gene expression can occur via several mechanisms that include

noncoding RNAs (e.g. microRNAs and long non-coding RNAs), biochemical modifications of DNA such DNA methylation or via histone protein modifications (Goldberg *et al*., 2007).

DNA methylation which can occur on cytosine bases in CpG rich sequences and especially at the promoter regions can lead to gene silencing through the inhibition of transcription factor binding and the changing of chromatin structure into a repressive state (Domcke *et al*., 2015). Many studies have shown that proper DNA methylation is essential in embryonic development (Bird, 2002; Santos *et al*., 2005; Bartolomei and Ferguson-Smith, 2011; Lomvardas and Maniatis, 2016). Research in our laboratory has identified an important role of *Hoxa2* in mouse palate development (Smith *et al*., 2009; Smith *et al*., 2013). In my study, I chose to investigate whether DNA methylation played a role in regulating the expression of *Hoxa2* in early mouse palate development. Three CpG islands were found to be situated in the *Hoxa2* promoter region and their methylation status did not change during the temporal expression of *Hoxa2* at any stages of palatal development. Further experiments were conducted to investigate the role of miRNA and long non-coding RNA in the regulation of *Hoxa2* expression.

microRNAs (miRNA) are ~22 nucleotide RNAs that can guide a RNA silencing pathway to regulate gene expression through their effect on messenger RNA (mRNA) levels (Usmani *et al*., 2016). The mature miRNAs interact with target mRNAs mostly within the 3' untranslated region (3' UTR) based on sequence complementarity between the miRNA and its target mRNA (Ha and Kim, 2014). Several *Hox* genes are regulated by miRNAs (Garaulet and Lai, 2015; Liu *et al*., 2013; Tang *et al*., 2014; Han *et al*., 2016); however, only one miRNA, miR-3960, has been identified to regulate *Hoxa2* gene expression in primary mouse osteoblasts (Hu *et al*., 2011). miR-669b has previously been found to be able to bind to the 3′UTR of insulin-like growth factor1 receptor (IGF1R) and down regulate its expression (Liang *et al*., 2011). I performed in

silico analysis and predicted six miRNAs that can potentially bind *Hoxa2* 3'UTR, namely: miR-376c, miR-669b, miR-431, miR-19a, miR-298 and miR-878-3p. Experimentally, I found miR-669b and miR-376c can both down regulate *Hoxa2* gene expression at the transcriptional and translational level. Moreover, two binding sites for miR-669b on *Hoxa2* 3′ UTR were identified using a luciferase reporter assay, while no direct binding of miR-376c was found on the *Hoxa2* 3′ UTR.

Hox genes are also regulated by epigenetic activators of the Trithorax group (TrxG) and epigenetic repressors of the Polycomb group (PcG) (Beck *et al*., 2010). The activation of *Hox* genes by TrxG can involve trimethylation of Histone 3 lysine 4 (H3K4me3) by mixed lineage leukemia 1 (MLL1)/ WD repeat domain 5 (WDR5) complex (Beck *et al*., 2010). Long noncoding RNAs (lncRNAs) have emerged with important regulatory roles in gene expression. LncRNAs *HOTTIP* and *Mistral* are both transcribed from *Hox* gene clusters and can introduce MLL1/WDR5 to nearby *Hox* gene promoters to induce H3K4me3 and activate gene expression (Wang *et al*., 2011, Bertani *et al*., 2011). *HOTAIRM1* is an antisense intergenic transcript transcribed between human *HOXA1* and *HOXA2* (Zhang *et al*., 2009). It can activate HOXA genes that are located at the 3' end of the HOXA cluster (3' HOXA genes) and its presence is necessary during myeloid differentiation. However, *HOTAIRM1* transcripts have not been reported in other species and little is known of how 3' HOXA genes are regulated by *HOTAIRM1*. In my study, I have identified a new transcript from the mouse *HoxA* cluster that shares sequence similarity with human *HOTAIRM1*, which we classify as the mouse *Hotairm1* (*mHotairm1*). I further demonstrated that *mHotairm1* can activate the expression of *Hoxa1* and *Hoxa2* by introducing MLL1/WDR5 to their promoters which in turn enhances H3K4me3 occupancy. Findings also indicate that WDR5 sumoylation could be important for its interaction with

mHotairm1. Taken together, my results provide additional evidence that connects lncRNAs to recruitment of epigenetic activators to promoters of target genes and for the first time provides a mechanism of the regulation of 3' *Hoxa* genes via *mHotairm1*.

CHAPTER 2

2. Literature Review

In the following literature review, I will introduce *Hox* genes and the roles of *Hoxa1* and *Hoxa2* in development and in disease. I will also provide an overview of epigenetics and its impact on gene expression and regulation. Epigenetic mechanisms including DNA methylation, common histone post-translational modifications, miRNAs and lncRNAs, as well as the crosstalk between these will be discussed. In my PhD thesis work, sumoylation of a WDR5 protein appeared to be important in the histone methylation mediated *Hox* gene expression that is directed by lncRNA, hence sumoylation is also included in this literature review.

2.1 *Hox* genes

Homeobox genes were first discovered in the fruit fly *Drosophila* (Bridges and Morgan, 1923), and are evolutionarily conserved transcription factors that regulate specific morphological patterns and cellular diversity along the anterior-posterior (AP) axis (Lewis, 1978; Pearson *et al*., 2005; Seifert *et al.,* 2015). Alterations in homeobox gene expression can result in the transformation of one body structure to resemble a homologous body structure in form and shape (Arlotta and Hobert, 2015; Tijchon *et al.,* 2015). This phenomenon was first observed in *Drosophila* and was termed ‘‘homeotic transformation’’ (Lewis, 1978; Hueber and Lohmann, 2008).

In *Drosophila*, a single homeotic complex (HOM-C) located on chromosome 3 consists of two separate clusters, Bithorax complex (BX-C) and Antennapedia complex (ANT-C) (Fig. 2.1) (Lewis, 1978). The antennapedia class of homeobox genes are generally referred to as

"*Hox"* genes. *Hox* genes are highly conserved and clustered into multigene loci (Lewis, 1978; Ruddle *et al*., 1994; Santini *et al*., 2003; Yu *et al*., 2012; De Kumar and Krumlauf, 2016). In vertebrates, *Hox* genes have duplicated and evolved to generate multiple sets of paralogous genes which are organized in four separate chromosomal clusters (McGinnis and Krumlauf, 1992; Bailey *et al*., 1997; Akin and Nazarali, 2005; Duboule, 2007; De Kumar and Krumlauf, 2016). They function to govern embryonic morphogenesis as well as cell differentiation (Pineault and Wellik, 2014; Seifert *et al.,* 2015). In the chromosomal cluster, all *Hox* genes are transcribed in the same 5' to 3' direction, and the more 3' a *Hox* gene is located on the chromosome, the more anteriorly it is expressed in the developing embryo (Gaunt *et al*., 1988; Favier and Dollé, 1997; Lappin *et al*., 2006; Mallo and Alonso, 2013).

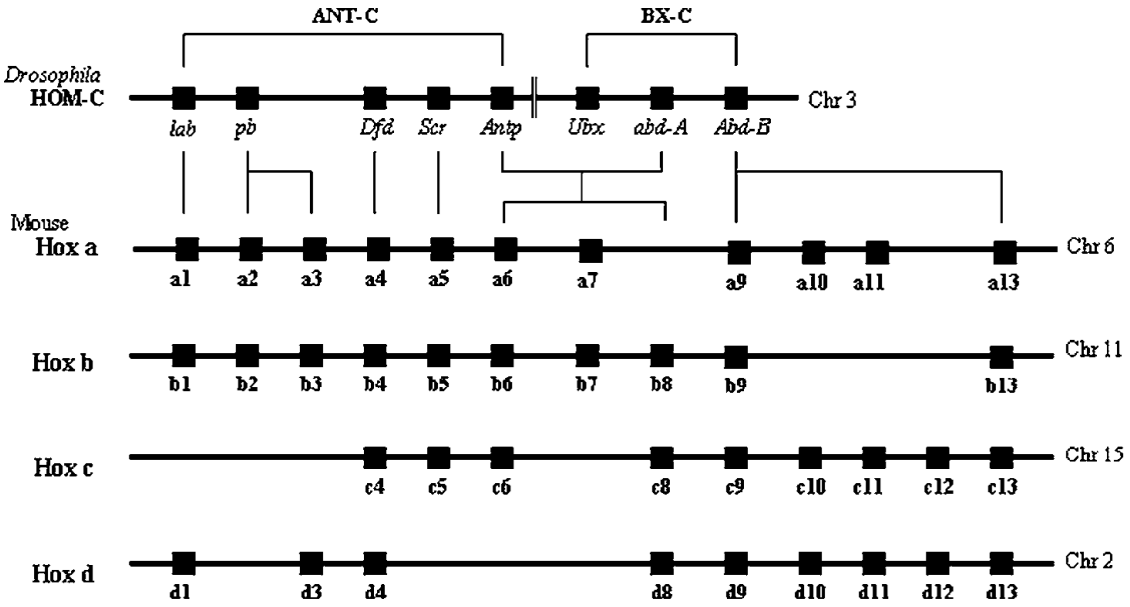


Figure 2.1. Schematic representation of *Drosophila* and *murine Hox* genes on chromosomal clusters. *Hox* genes are illustrated by black boxes. There are two clusters in the *Drosophila* HOM-C: Antennapedia (ANT-C) and Bithorax (BX-C). ANT-C is composed of: *lab = labial*, *pb = proboscipedia*, *Dfd = Deformed*, *Scr = Sex Combs Reduced*, and *Antp = Antennapedia*; BX-C is composed of: *Ubx = Ultrabithorax*, *abd-A = abdominal A*, *Abd-B = Abdominal-B*. There are 39 *murine Hox* genes present on four separate (Hox a, Hox b, Hox c, Hox d) chromosomal clusters. This figure is taken from Akin and Nazarali (2005) with permission.

Hox genes are very important as they regulate numerous pathways in embryonic development and other biological processes (Fig. 2.2). All homeobox genes contain a "homeodomain" which is defined as a class of protein domain that contain a conserved 60 amino acid region encoded by 180 bp homeobox sequence (Scott *et al*., 1989; Nazarali *et al*., 1992; Banerjee-Basu and Baxevanis, 2001; Anderson *et al*., 2012; Bürglin and Affolter, 2016). Hox proteins can enhance or suppress the expression of downstream genes through the affinity between its homeodomain and target DNA sequences (McGinnis and Krumlauf, 1992; Gehring, 1993; Kumar and Nazarali, 2001; Shah and Sukumar, 2010). The core structure of this domain consists of four alpha helices (Kissinger *et al*., 1990; Gehring *et al*., 1994; Akin and Nazarali, 2005; Bürglin and Affolter, 2016). The coordinate activity between the N-terminal region of the first helix and cofactors within a transcriptional complex is required for Hox proteins to control specific gene expression and affect cell division and cellular functions (Mann and Affolter, 1998; Akin and Nazarali, 2005; Mann *et al*., 2009). The second and third helix form an evolutionarily conserved helix-turn-helix motif that is responsible for the recognition and binding of target DNA sequences (Gehring *et al*., 1994; Akin and Nazarali, 2005; Bürglin and Affolter, 2016). The hydrophilic face of helix 3 contacts the target DNA sequence in the major groove (Kissinger *et al*., 1990; Banerjee-Basu and Baxevanis, 2001; Svingen and Tonissen, 2006; Bürglin and Affolter, 2016). In *vitro* analysis determined an element of approximately 10-12 bases with a core sequence of TAAT to be a core binding site for Hox proteins (Kissinger *et al*., 1990; Kumar and Nazarali, 2001; Akin and Nazarali, 2005; Breitinger *et al*., 2012; Beh *et al*., 2016).

In mice, there are four separate chromosomal clusters (Hox A, B, C, and D) composed of 39 genes in total (Fig. 2.1). These *Hox* genes are arranged in a 3′ to 5′ order in each cluster with synteny to the *Drosophila* HOM-C (Fig. 2.1), with sequence similarity between genes on

different clusters, as well as their positions on the chromosomes (Akin and Nazarali, 2005; Mallo and Alonso, 2013).

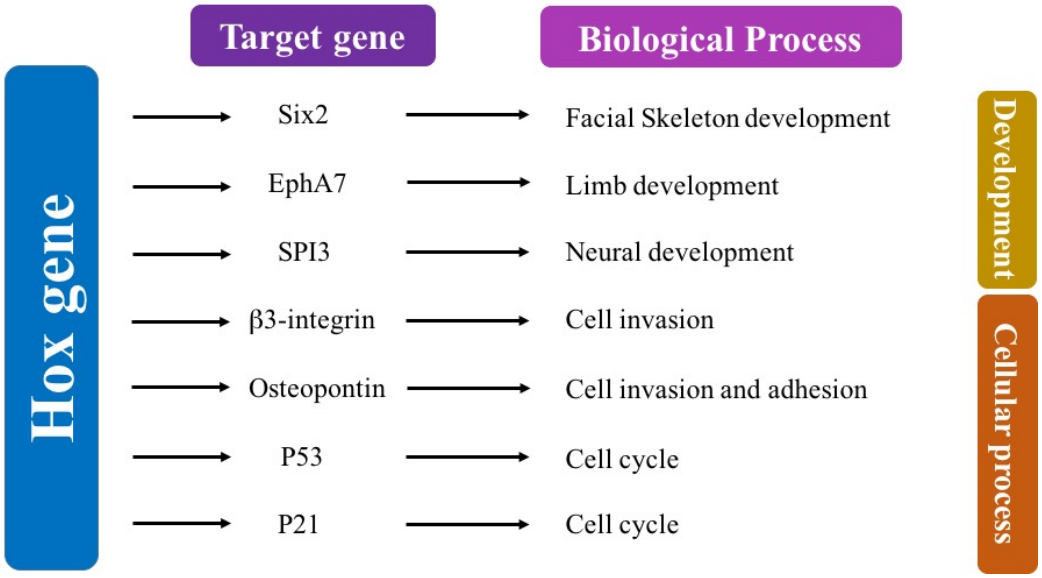


Figure 2.2. Examples of biological pathways that are regulated by Hox genes through their regulation of downstream targets. Several Hox-targeted genes that are important in development and cellular processes are shown. This figure is modified from Svingen and Tonissen (2006).

2.1.1 *Hoxa1* gene in embryonic development

In embryonic development, *Hoxa1* is one of the first *Hox* genes to be activated (Murphy and Hill, 1991). In mouse, starting from embryo day 7.5 (E7.5), *Hoxa1* expression can be detected in the primitive streak, in newly formed mesoderm, and overlying neuroectoderm (Fig. 2.3; Murphy and Hill, 1991). *Hoxa1* expression extends anteriorly, and between E7.75 and E8.25 its expression reaches the most anterior border in the presumptive hindbrain where its expression lasts for aprroximately 12 h (Murphy and Hill, 1991). By E8.5, *Hoxa1* has retreated from this region and remains in more posterior regions (Murphy and Hill, 1991; Makki and Capecchi, 2010). This expression pattern suggested a role of *Hoxa1* in hindbrain development. In fact,

Hoxa1 knockout mice exhibit abnormal hindbrain development (Studer *et al*.; 1998; Helmbacher *et al*., 1998; Tischfield *et al*., 2005) and die shortly after birth from breathing defects (Chisaka *et al*., 1992). *Hoxa1* null embryonic stem (ES) cells lack the ability to differentiate along a neural cell lineage following retinoic acid induction and express lower levels of neuronal differentiation markers compared to normal wild-type ES cells (Martinez-Ceballos and Gudas, 2008). *Hoxa1* acts upstream of four genes involved in hindbrain and neuron development: hepatocyte nuclear factor 1-beta (*Hnf1b*), forkhead box D3 (*Foxd3*), Zic family member 1 (*Zic1*) and LIM homeobox 5 (*Lhx5*) (Makki and Capecchi, 2011). *Hoxa1* has also been proposed as a candidate gene for autism spectrum disorder (ASD), a common neurodevelopmental condition in children and adolescents (Song *et al*., 2011; Raznahan *et al*., 2012). A specific *HOXA1*-A218G mutation has received particular attention in ASD since the A218G genotype has the capacity to modify the rate of cerebellar growth (Raznahan *et al*., 2012).

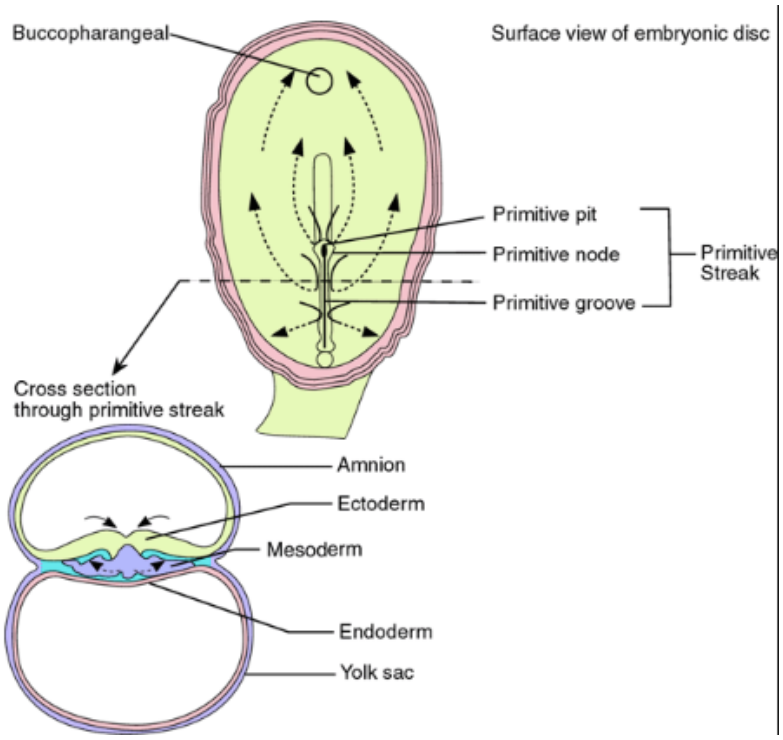


Figure 2.3. Schematic diagram of primitive streak and germ layers in mouse embryo. (https://o.quizlet.com/9elYyZW.TOVAN103uQlAqw_m.png)

In addition to the hindbrain, *Hoxa1* also affects inner ear development (Makki and Capecchi, 2010). Although researchers previously believed that *Hoxa1* expression was absent in mouse inner ear and the effects of loss of *Hoxa1* on inner ear development was indirect (Murphy and Hill, 1991), Makki and Capecchi (2010) reported significant *Hoxa1* cell lineage expression in the developing mouse otic epithelium, which raised the possibility that *Hoxa1* may have a direct role in inner ear patterning. Several genes regulating ear development have also been identified as downstream targets of *Hoxa1*. In *Hoxa1* null mice, paired box 8 (*Pax8*) and fibroblast growth factor receptor 3 (*Fgfr3*), known to be important in ear development, were identified as downstream targets of *Hoxa1* (Makki and Capecchi, 2011). A mutation in the *Hoxa1* coding sequence in pigs resulting in a truncated protein lacking the homeodomain also induced malformations of both the outer and middle ears. At the genetic level, this mutation of *Hoxa1* affected the expression of fibroblast growth factor 1 (*FGF1*) and *FGFR3* in the FGF signaling pathway, a pathway that is crucial in ear pinna development (Qiao *et al.,* 2015).

Researchers have also demonstrated a previously unrecognized role of *HOXA1* in cardiovascular development in humans (Tischfield *et al*., 2005). This finding lends credence to Makki and Capecchi's findings that show in mice all cardiac neural crest cells in the outflow tract are derived from *Hoxa1*-expressing cells (Makki and Capecchi, 2011). Further research demonstrated that *Hoxa1* knockout mice exhibit severe cardiovascular, cerebrovascular and glandular defects (Makki and Capecchi, 2012). *Hoxa1* expression in cardiac progenitor cells (Bertrand *et al*., 2011) is known to form a gene regulatory network with Hoxb1 in the formation of the cardiac outflow tract (Roux *et al*., 2015). At the molecular level, Hoxa1 regulates the expression of *Hnf1b, Foxd3* and *Zic1* that are necessary for neural crest specification (Makki and Capecchi, 2011; Makki and Capecchi, 2012). In humans, homozygous truncation mutations in

HOXA1 can lead to Bosley–Salih–Alorainy syndrome (BSAS) or Athabascan brainstem dysgenesis syndrome (ABDS), from which patients can exhibit facial weakness, hypoventilation, swallowing dysfunction, deafness and heart defects (Bosley *et al*., 2007; Holve *et al*., 2003; Jin and Sukumar, 2016).

2.1.2 *Hoxa2* gene in embryonic development

Hoxa2 is a member of the *Hox* gene family that shares a 60 amino acid homeodomain and encodes a 41-kDa protein (Nazarali *et al*., 1992; Tan *et al*., 1992; Smith *et al*., 2009). The *Hoxa2* protein binds to downstream genes to regulate their expression in a spatio-temporal manner (Akin and Nazarali, 2005). *Hoxa2* gene acts as the selector gene for second branchial arch patterning (Gendron-Maguire *et al*., 1993; Minoux *et al*., 2013; Cox *et al*., 2014) and plays an essential role in early embryo development.

Hoxa2 is important for the development of skeletal structures derived from the second pharyngeal arch (PA) (Fig. 2.4A) and it determines the areas of skeletogenesis from the second PA mesenchyme by an inhibitory mechanism (Kanzler *et al*., 1998). In *Hoxa2* knockout mice, the structures in the second PA are arranged as a mirror image of the first PA (Fig. 2.4A), indicating *Hoxa2* may prevent the formation of the first PA (Gendron-Maguire *et al*., 1993). Tavella and Bobola (2009) have shown that the over expression of *Hoxa2* in mice can cause failure of bones to form in the cranial base. During chondrogenesis, overexpression of *Hoxa2* in cells entering chondrogenesis can impair cartilage development and lead to embryonic delay of ossification followed by a postnatal proportionate short stature with reduction in the length of the trunk and limbs (Massip *et al*., 2007; Deprez *et al*., 2012). At the molecular level, the persistent expression of *Hoxa2* downregulates expression of genes controlling cell differentiation in

chondrogenesis, namely bone morphogenetic protein 7 (Bmp7), msh homeobox 1 (*Msx1*), paired box 9 (*Pax9*), sex determining region Y- box 6 (*Sox6*), *Sox9* and Wnt family member 5a (*Wnt5a*) (Deprez *et al*., 2013). Ectopic expression of *Hoxa2* in the *Hox*-negative cranial neural crest cells (CNCCs) in mice resulted in skeletal defects including absent or reduced bones in the skull vault and maxillary structures (Kitazawa *et al*., 2015).

Hoxa2 also plays a role in ear development where a loss of *Hoxa2* function in mice can cause an abnormal appearance of the external ear, a condition that is known as microtia (Kanzler *et al*., 1998; Santagati *et al*., 2005). Several bones in the middle ear are affected in *Hoxa2*$^{-/-}$ mice, including stapes, the styloid bone and the lesser horn of the hyoid bone (Fig. 2.4B; Kanzler *et al*., 1998; Minoux and Rijli, 2010). In the mouse the whole auricle is derived from the *Hoxa2*-expressing second PA neural crest-derived mesenchyme and *Hoxa2* spatially organises cell proliferation during external ear development (Minoux *et al*., 2013). Further genetic analyses showed that *Hoxa2* regulates the expression of bone morphogenetic protein 4 (*Bmp4*), bone morphogenetic protein 5 (*Bmp5*), twisted gastrulation (*Tsg*) (Minoux *et al*., 2013; Cox *et al*., 2014) and H6 family homeobox 1 (*Hmx1*) (Rosin *et al*., 2016) in the developing pinna. In human studies, several mutations of HOXA2 coding region have been related to microtia (Alasti *et al*., 2008; Monks *et al*., 2010). In some cases, HOXA2 haploinsufficiency (with only one functional copy of HOXA2 gene and the other copy inactivated by mutation) is also known to cause microtia and hearing loss (Brown *et al*., 2013; Jin and Sukumar, 2016).

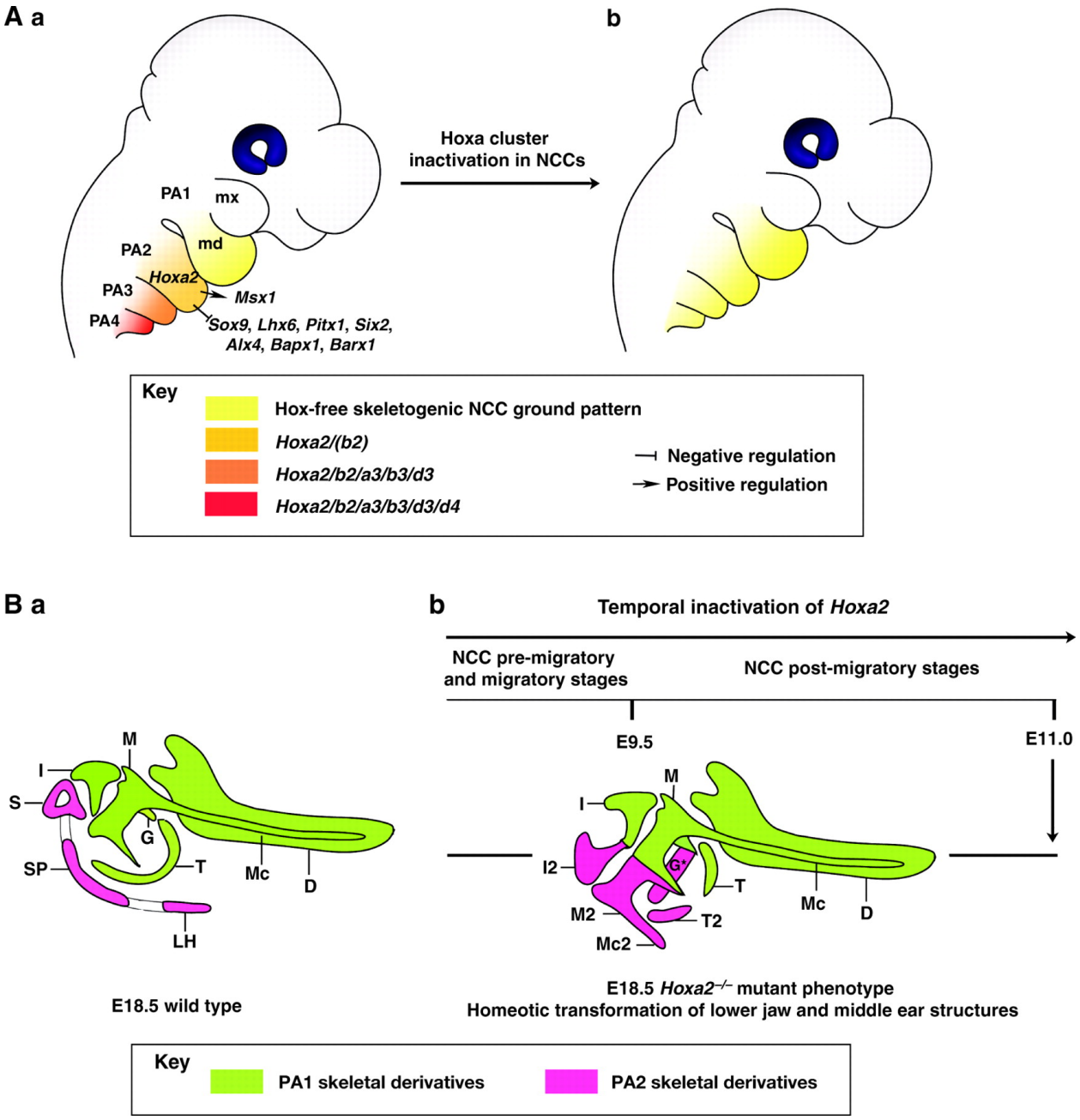


Figure 2.4. Schematic diagram of mouse pharyngeal arches (PA) and developing ear. A. Schematic of E10.5 mouse embryo was shown with specific Hox code in each PA (a). Inactivation of Hoxa genes in NCCs resulted in a Hox-free ground pattern in all four PAs (b). B. Schematic diagram of the effect of *Hoxa2* on PA2 derived lower jaw and middle ear structures. D – dentary bone; G – gonial bone; G* – modified gonial bone; I and I2 – incus and its duplicated counterpart; LH – lesser horns of the hyoid bone; M and M2 – malleus and its duplicated counterpart; Mc and Mc2 – Meckel's cartilage and its duplicated counterpart; SP – styloid process; S – stapes; T and T2 – tympanic bone and its duplicated counterpart. Figure taked from Minoux and Rijli (2010) with permission.

A previous study from our laboratory has shown that *Hoxa2* plays a direct role in murine palate development (Smith *et al*., 2009). In mice at embryonic day 11.5 (E11.5), the CNCC-derived cells bud out of the maxillary prominences to form two downward projecting structures of the palatal shelves. These shelves then undergo significant growth and grow vertically down the sides of the tongue until E13.5. At this point the two shelves begin to ascend and grow horizontally above the tongue. At E14.5, the palatal shelves contact one another and fuse completely at E15.5 (Ferguson, 1988) (Fig. 2.5). *In vivo Hoxa2* protein is expressed in both epithelial and mesenchymal cells of the developing murine secondary palate (Nazarali *et al*., 2000). Our laboratory has further demonstrated a spatial and temporal expression of *Hoxa2* mRNA and protein within the mouse palate from E12.5 to E15.5 (Smith *et al*., 2009). During development, palatal shelves emerge, elevate, and then fuse. The *Hoxa2* expression pattern changes during this period as well, being highly expressed at the beginning of palatal development, peaking at E13.5 and declining significantly after E14.5 (Smith *et al*., 2009) (Fig. 2.5). Increased cell proliferation and decreased fusion rates were also observed in *Hoxa2* null mice. Several genes involved in palate development, including *Msx1*, *Bmp4*, *Barx1* and *Ptx1*, have been identified as downstream targets of *Hoxa2* within the palate (Smith *et al*., 2009).

During nervous system development, *Hoxa2* expression affects development of hindbrain and cerebellum through its impact on rhombomeres (Gavalas *et al*., 1997). In the vertebrate central nervous system (CNS), oligodendrocytes form myelin sheath. Our laboratory has shown that *Hoxa2* is expressed throughout oligodendrogenesis (Nicolay *et al*., 2004). In a mouse model of *Hoxa2* loss-of-function, early stages of oligodendrogenesis does not appear to be altered in the spinal cord (Nicolay *et al*., 2004), although over expression of *Hoxa2* inhibits oligodendrogenesis throughout the brain (Miguez *et al*., 2012). In addition, we have shown that

overexpression of *Hoxa2* in CG4 oligodendroglial cells increased cell proliferation but delayed CG4 oligodendroglial cell differentiation (Wang *et al*., 2011).

As summarized above, *Hoxa1* and *Hoxa2* genes control many important embryonic developmental processes, hence it would be of great value to investigate how expression of *Hoxa1* and *Hoxa2* are regulated, especially at the epigenetic level.

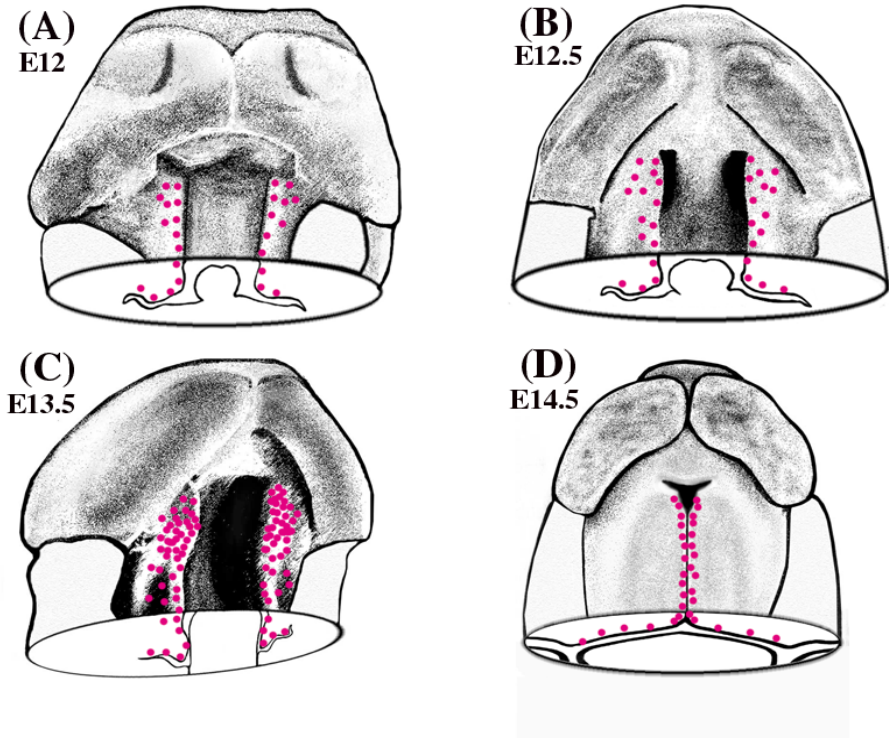


Figure 2.5. Mouse palate development and *Hoxa2* expression. The developing secondary palate first extends vertically down the sides of the tongue from E11.5-E12.5 (A and B). This is followed by the two palatal shelves ascending horizontally above the tongue at E13.5 (C) and extending towards each other at E14.5 (D). The red dots in the figure indicate the expression of *Hoxa2* at different stages of mouse palate development. The density of the dots represents the relative expression of *Hoxa2*. Figure taken from W. Zhang's M.Sc., Thesis 2003.

2.1.3 Hox gene activation by all-*trans* retinoic acid (ATRA)

All-*trans* retinoic acid (ATRA) is derived from vitamin A. It can activate the early expression of 3' *Hox* paralogs and is important in embryo development (Daftary and Taylor, 2006; Cunningham and Duester, 2015). ATRA can activate gene expression via its cognate receptors, including retinoic acid receptors (RARs) and retinoid X receptors (RXRs). RAR and RXR can form a heterodimer complex and bind to target gene at a retinoic acid response element (RARE) located in enhancer regions (Cunningham and Duester, 2015). Without activation of RA binding, unliganded RAR-RAX heterodimer bind to RARE sequences to maintain repressed transcription. The unliganded RAR-RAX heterodimer can also recruit co-repressors to further help maintain repressed transcription. Upon the binding of ATRA to RAR-RXR heterodimer, a conformational change is induced in the heterodimer which leads to the recruitment of gene co-activators and further induces gene expression (Daftary and Taylor, 2006; Cunningham and Duester, 2015). RARE sequences are found in many *Hox* genes and ATRA induces sequential activation of *Hox* genes that correlates with their positions on the chromosomal cluster (Kashyap et al., 2011; Cunningham and Duester, 2015). Following ATRA exposure, epigenomic reorganization of the Hox cluster occurs quickly (Kashyap et al., 2011). The 3'-Hoxa1 RARE is essential for ATRA mediated activation of Hoxa and Hoxb clusters (Kashyap et al., 2011). Important roles of ATRA have been reported in development including central nervous system development (Daftary and Taylor, 2006), limb development and organ development (Cunningham and Duester, 2015).

2.2 Epigenetics

Epigenetics refers to the heritable changes in gene expression without any changes in DNA sequences and is a major mechanism that regulates gene expression changes in response to gene-environment interactions (Holliday, 2006). At the molecular level, epigenetic pathways include regulation by noncoding RNAs and biochemical modifications of the DNA and histone proteins, such as methylation, acetylation and phosphorylation (Skinner, 2011; Yao *et al.,* 2016). The importance of epigenetic modifications has long been recognized in the areas such as stem cell research, cancer (Pogribny, 2010) and developmental biology (Skinner, 2011; Yao *et al.,* 2016).

2.2.1 DNA methylation

DNA methylation is one of the most widely studied aspects in epigenetic mechanism. Proper DNA methylation is essential for mammalian embryonic development and is involved in gene repression, regulation of parental imprinting and X-chromosome inactivation (Bird, 2002; Santos *et al*., 2005; Bartolomei and Ferguson-Smith, 2011; Lomvardas and Maniatis, 2016). It is also closely related to diseases like cancer, neurodegenerative diseases, psychiatric disorders and cardiovascular diseases, and is a potential therapeutic target in many diseases (Abdelfatah *et al*., 2016; Tang *et al*., 2016; Li *et al*., 2016; Fries *et al*., 2016; Napoli *et al*., 2016; Zhong *et al*., 2016). DNA methylation occurs on cytosine bases throughout the genome but is most relevant when present in sequences rich in CpG dinucleotides, which are called CpG islands, often found in promoter regions. Methylated DNA contains the covalent addition of a methyl group to cytosine residues at CpG dinucleotides. The enzymes that catalyze this reaction are the DNA methyltransferases (DNMTs). Mammalian DNMTs are comprised of three regions: (1) an N-

terminal regulatory region which guides the localization of the enzymes to the nucleus and mediates their interactions with other proteins, DNA and chromatin; (2) a central linker region; and (3) a C terminal catalytic region which catalyzes the transfer of the methyl group from a cofactor molecule S-adenosyl-l-methionine (AdoMet or SAM) to the C5 position of the cytosine residue (Fig. 2.6; Jurkowska *et al*., 2011; Uysal *et al*., 2015).

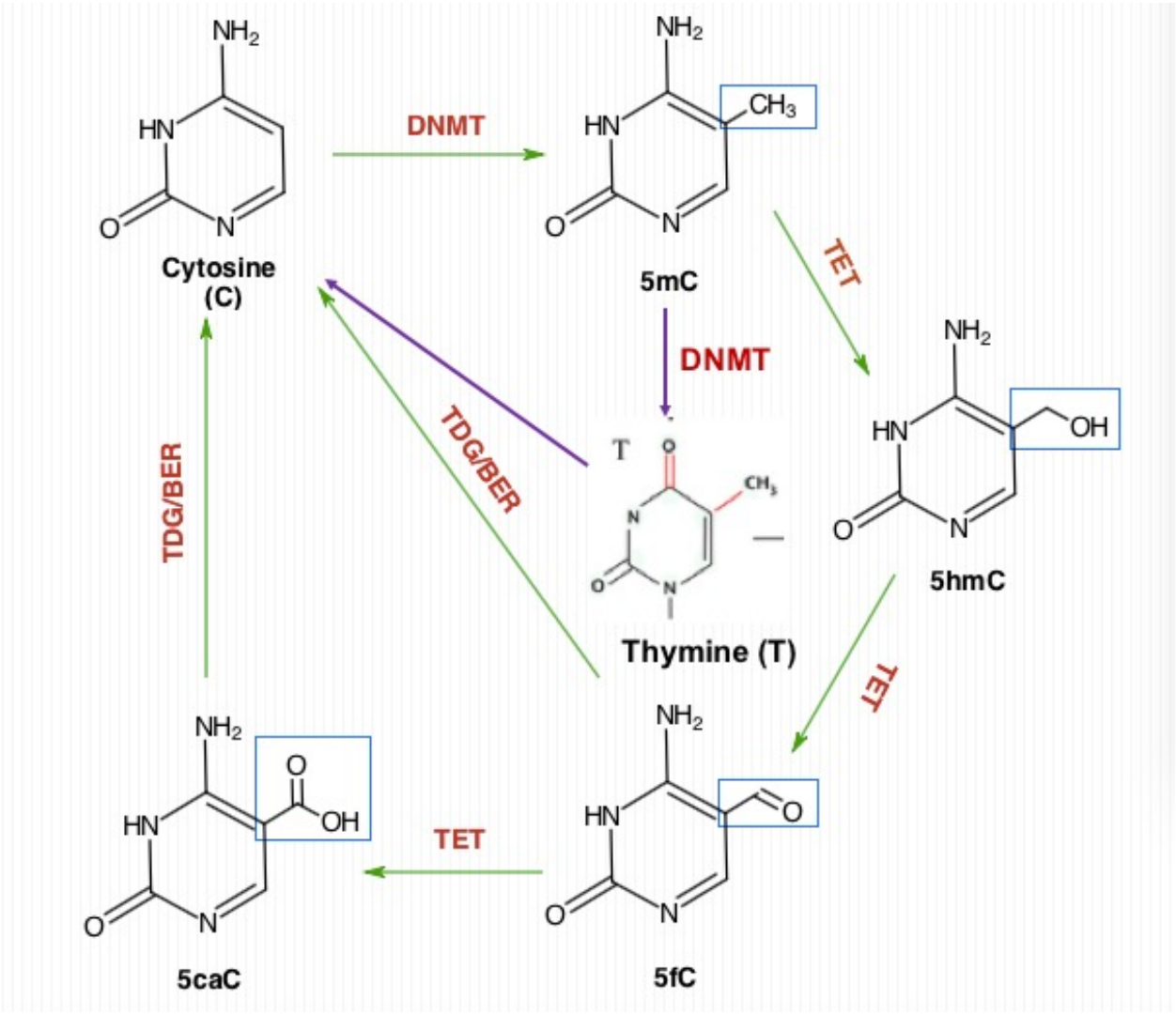


Figure 2.6. Schematic diagram of cytosine's methylation and demethylation processes. DNA methylation on 5-cytosine (5mC) is carried out by DNMTs. 5mC can be demethylated by the ten-eleven translocation (TET) familly of DNA hydroxylases and converted back to cytosine through thymine DNA glycosylase (TDG)/ base excision and repair (BER) mechanism. 5mC can also be deaminiated and converted into thymine (T) and further be replaced by unmethylated C through BER mechanism. Figure modified from Xu *et al.,* 2014.

Table 2.1. List of structures and functions of DNMTs

DNMTs	Structure			Function	Reference
	N-terminal regulatory region	Linker region	C terminal catalytic region		
DNMT1	√	√	√	Copy pre-existing methylation patterns from parental strand onto newly synthesized daughter strand.	Bestor, 1992; Uysal *et al*., 2015
DNMT2		√	√	Transfer RNA methylation.	Goll *et al*., 2006; Jeltsch *et al*., 2016
DNMT3a	√	√	√	Introduce new DNA methylation patterns in mammalian development and in germ cells.	Okano *et al*., 1999; Uysal *et al*., 2015
DNMT3b	√	√	√		
DNMT3L	√	√		Facilitate DNMT3a and DNMT3b in DNA methylation.	Uysal *et al*., 2015; Basu *et al*., 2016

Table 2.2 List of proteins that bind to methylated DNA

Methyl-DNA binding Domain	Major group members	Binding preference	Function	References
MBDs	MeCP2, MBD1, MBD2	Single 5mC	Transcriptional repressors; interact with gene co-repressors	Du *et al*., 2015; Wood and Zhou, 2016
	MBD3	5mC, 5hmC, C	Transcriptional repressors	Shimbo *et al*., 2013; Du *et al*., 2015
	MBD4	Single 5mC	DNA damage repair	Laget *et al*., 2014
The methyl-CpG binding zinc-finger proteins	Kaiso	Two methyl-CpG	Transcriptional repressors	Sasai *et al*., 2010; Bogdanović and Veenstra, 2009
	ZBTB4, ZBTB38	Single 5mC	Transcriptional repressors	
SET and SAR family proteins	UHRF1	Hemi-methylated DNA	Associate with DNMT1; DNA damage repair; recruitment of KDAC	Unoki *et al*., 2004; Berkyurek *et al*., 2014; Yang *et al*., 2013
	UHFR2	5hmC	DNA damage repair	Liu *et al*., 2016b; Yang *et al*., 2013

In mammals, five different DNMTs, namely DNMT1, DNMT2, DNMT3a, DNMT3b and DNMT3L have been identified and amongst all five DNMTs, the three active DNMTs are DNMT1, DNMT3a and DNMT3b (Table 2.1; Uysal *et al*., 2015). During DNA replication, hemimethylated DNA is created with only the parental strand having methylation marks while the newly synthesized strand remains unmethylated. DNMT1 preferentially recognizes hemimethylated DNA over unmethylated DNA and can copy pre-existing methylation patterns onto the new DNA strand to remethylate the daughter strand (Bestor, 1992; Uysal *et al*., 2015). In this way DNA methylation can be preserved during DNA replication. The *de novo* methyltransferases DNMT3a and DNMT3b are mainly responsible for establishing DNA methylation patterns in mammalian development and in germ cells by introducing cytosine methylation to previously unmethylated CpG sites (Okano *et al*., 1999). Unlike DNMT1, they do not show any significant preference between hemimethylated and unmethylated DNA (Okano *et al*., 1998; Uysal *et al*., 2015). Besides methylation of CpG sites, DNMT3a can also methylate cytosine residues at non-CpG sites but the biological function of these methylation sites remain unknown (Arand *et al*., 2012; Shirane *et al*., 2013). DNMT3s are also able to read other epigenetic marks, like trimethylation of histone 3 lysine 4 (H3K36me3), to guide DNA methylation (Rondelet *et al*., 2016). In recent years, a function in the DNA demethylation pathway has also been identified for DNMT3a and DNMT3b. It has been reported that under low SAM concentrations, DNMT3a and DNMT3b can convert 5-methylcytosine (5mC) into T, which could further be replaced by an unmodified C by the base excision and repair (BER) mechanism (Chen *et al*., 2013; van der Wijst *et al*., 2015). DNMT3L does not contain methyltransferase active site motifs and thus lack DNA methyltransferase activity but it is also present in mammals and is functionally related to DNMT3a and DNMT3b and modulates their

catalytic activity (Suetake *et al*., 2004; Uysal *et al*., 2015). It can bind to the N terminus of Histone 3 to bring about DNA methylation and the recognition of unmethylated H3K4 by DNMT3L is important for the methylation function of DNMT3a (Basu *et al*., 2016; Hu *et al*., 2009; Ooi *et al*., 2007). Due to a lack of an N-terminal regulatory domain, DNMT2 cannot catalyze DNA methylation. However, a functional study of DNMT2 showed that it is involved in methylation of transfer RNA through its C-terminal catalytic domain (Goll *et al*., 2006; Jeltsch *et al*., 2016).

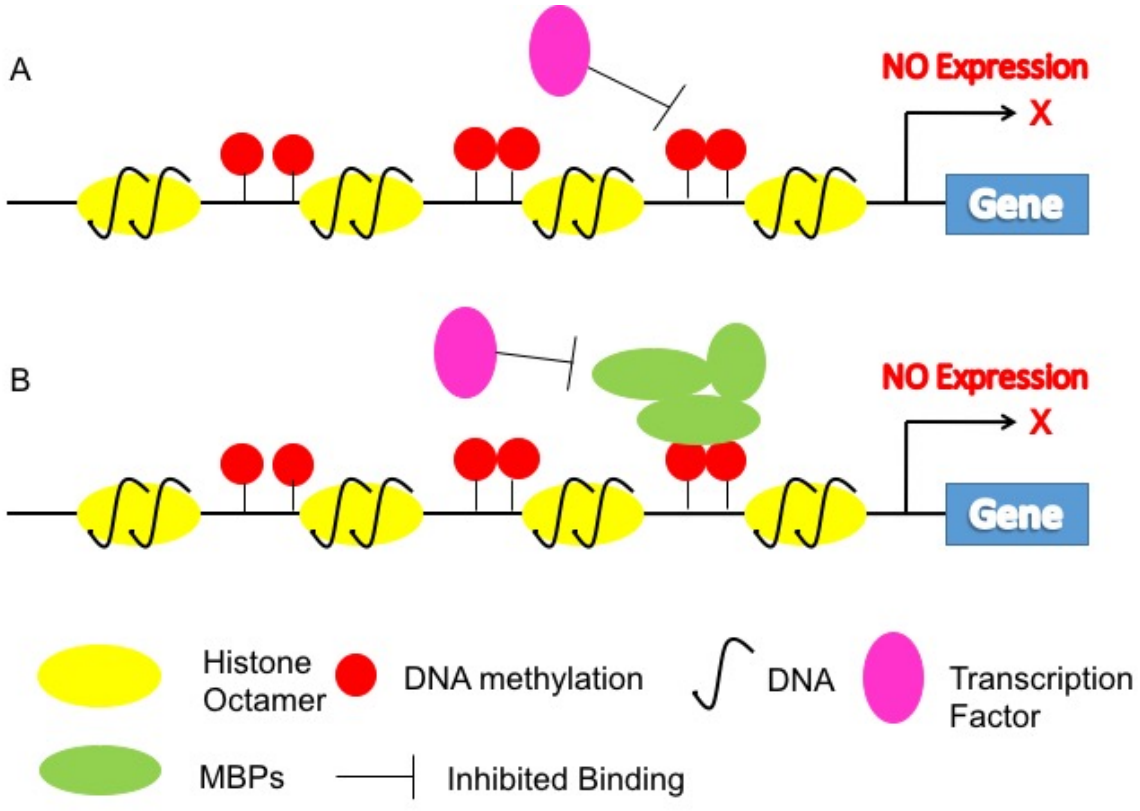


Figure 2.7. Mechanisms of DNA-methylation-mediated gene repression. (a) DNA methylation directly inhibit the binding of some transcription factors (TF). (b) Methyl-CpG-binding proteins (MBPs) can recognize methylated DNA and recruit co-repressor molecules to silence transcription. They can also modify surrounding chromatin structure to achieve gene silencing.

DNA methylation usually occurs at the transcription start site and generally prevents transcription in two ways (Fig 2.7). First it can directly inhibit the binding of transcription factors or regulators to the DNA sequence (Fig 2.7a). Secondly, it can also indirectly recruit methyl-CpG binding proteins (MBPs) which bind to and recognize 5-methylcytosines, to change chromatin into a repressive state (Fig 2.7b) (Bogdanovic and Veenstra, 2009; Deaton and Bird, 2011; Moore *et al*., 2013; Domcke *et al*., 2015). DNA methylation can also occur within the coding region of the gene and was initially believed to inhibit transcriptional elongation (Klose and Bird, 2006) although later evidences showed that this methylation was not associated with transcriptional repression (Jones, 2012). However, it has been demonstrated that gene body methylation is positively correlated with active transcription, and further research has suggested that it may play a role in gene splicing (Shukla *et al*., 2011; Jones, 2012). The function of methylation within the coding region however, needs further investigation (Hellman and Chess, 2007; Jones, 2012; Yang *et al*., 2014). Based on their functional domains, there are three families of proteins that bind to methylated DNA (Table 2.2), namely; (1) Methyl-CpG binding domain proteins (MBDs) including MeCP2, MBDs 1-6, SETDB1, SETDB2, BAZ2A and BAZ2B (reviewed by Fan and Hutnick, 2005; Moore *et al*., 2013; Du *et al*., 2015); (2) the methyl-CpG binding zinc-finger proteins of the Kaiso family where so far only three members have been described; Kaiso, ZBTB4 and ZBTB38 (Prokhortchouk *et al*., 2001; Sansom *et al*., 2007; Moore *et al*., 2013; Du *et al*., 2015); and finally (3) SET (Suppressor of variegation-enhancer of zeste-trithorax) and the ring finger-associated (SAR) family composed of UHRF1 and UHFR2 (Pichler *et al*., 2011; Moore *et al*., 2013; Du *et al*., 2015). Due to amino acid changes at critical positions in the methyl-CpG binding domain, MBD3 is the only MBD protein that binds to methyl-CpG, 5-hydroxymethylated DNA (5hmC) as well as unmethylated DNA (Saito and Ishikawa, 2002;

Yildirim *et al*., 2011; Shimbo *et al*., 2013). MBD4 can also act as a DNA repair enzyme (Walsh and Xu, 2006; Laget *et al*., 2014). MeCP2, MBD1 and MBD2 are transcriptional repressors and these MBD proteins were found to interact with co-repressors such as histone deacetylases, methyltransferases, and chromatin remodeling factors that regulate gene expression by acting on histone proteins (Sasai and Defossez, 2009; Du *et al*., 2015; Wood and Zhou, 2016). Similar to MBDs, ZBTB4 and ZBTB38 require only one methylated CpG for binding, while Kaiso requires two methyl-CpGs for binding (Sasai *et al*., 2010; Bogdanović and Veenstra, 2009). UHRF1 acts as gene repressor via recruitment of the KDAC complex to the promoters (Unoki *et al*., 2004). UHRF1 is associated with DNMT1 as it also has a preferential binding affinity for hemi-methylated DNA and specifically recruits DNMT1 to hemi-methylated sites during DNA replication (Bostick *et al*., 2007; Berkyurek *et al*., 2014). UHRF1 may also be important in histone modification (Sasai and Defossez, 2009; Qin *et al*., 2015). UHRF2 is a paralog of UHRF1 but have structural differences at the SRA domains. Due to this structural difference, UHRF2 has higher affinity to 5-hydroxymethylated cytosine (Liu *et al*., 2016b). UHRF1 and UHRF2 are also involved in DNA damage repair (Luo *et al*., 2013; Yang *et al*., 2013). Researchers also find DNA methylation in actively transcribed genes, thus it may also play a positive role in transcription regulation (Weber *et al*., 2007).

DNA methylation is especially important in mammalian embryo development. Mutation in mouse *Dnmt1* gene causes extensive demethylation of the genome and leads to embryonic lethality shortly after gastrulation (Li *et al*., 1992; Kurihara *et al*., 2008; Arand *et al*., 2012). *Dnmt3b*$^{-/-}$ mice have multiple developmental defects and embryos die at embryonic day E9.5 while *Dnmt3a*$^{-/-}$ mice die shortly after birth (Okano *et al*., 1999). Loss of *Dnmt3a* in mouse embryonic cardiomyocytes alteres multiple signaling pathway in cardiomyocytes and inhibits

function of embryonic cardiomyocytes (Fang *et al*., 2016). DNMTs are necessary for bovine parthenogenetic preimplanatation embryo development as DNMT inhibitors can inhibit this process (Zhang *et al*., 2015). Unlike knockout of DNMTs, single knockout of MBDs in mice display only mild changes. This could be because MBDs can compensate each other or the silencing of genes by DNA methylation can also occur through different pathways (Sasai and Defossez, 2009; Du *et al*., 2015). MeCP2 is related to neural development as mutation of MeCP2 can lead to severe neurodevelopmental disorders in females (Amir *et al*., 1999; Yang *et al*., 2016). Similar to DNMT1 mutants, UHRF1 knockout mice die during early embryogenesis (Sharif *et al*., 2007). Hence all this evidence suggest that proper DNA methylation status is critical for embryonic development. As reviewed above, *Hox* genes are also essential in embryonic development. Thus it would be interesting to investigate how expression of *Hox* genes is controlled by DNA methylation during embryonic development.

2.2.2 Histone modification

Histone modifications can lead to either the activation or the silencing of gene transcription. Posttranslational histone modifications normally occur on the N-terminal tails of Histone H3 and H4. There are many types of modifications of histone proteins, including acetylation, methylation, phosphorylation, ubiquitination, sumoylation and ADP-ribosylation (Jenuwein and Allis, 2001; Spivakov and Fisher, 2007; Bannister and Kouzarides, 2011; Canovas and Ross, 2016). Histone modifications can regulate gene expression in three ways. First, they can regulate chromatin structure to alter DNA accessibility (Abel and Zukin, 2008; Canovas and Ross, 2016). Second, they serve as a signal by integrating responses to multiple biochemical signaling cascades and recruit the transcriptional machinery and chromatin

remodeling complexes (Abel and Zukin, 2008; Badeaux and Shi, 2013). Third, histone modifications can also mediate epigenetic changes in gene expression (Abel and Zukin, 2008; Moore *et al*., 2013). Below, I review three common histone modifications: histone acetylation, histone phosphorylation and histone methylation.

2.2.2.1 Histone acetylation

Histone acetylation occurs on a lysine residue in histones H2A, H2B, H3 and H4 and is associated with transcriptional activation by opening the chromatin structure (Graves *et al*., 2016). Histone acetylation not only takes place on the N-terminal tail of histone proteins, but also to a lesser extent in the globular histone core (Bannister and Kouzarides, 2011; Graves *et al*., 2016). Histone acetylation of lysine is regulated by lysine acetyltransferases (KATs, also known as histone acetyltransferases, HATs) and lysine deacetylases (KDACs, also known as histone deacetylases, HDACs). The KATs can transfer an acetyl group to the lysine side chain, which neutralizes the lysine's positive charge to weaken the electrostatic interactions between histones and the DNA phosphodiester backbone (Bannister and Kouzarides, 2011; Canovas and Ross, 2016) (Fig 2.8). Based on the structural and functional similarity of their catalytic domains, KATs are grouped into five families: p300/CBP family, the MYST-family, the GCN5-related N-acetyltransferase (GNAT) superfamily, the nuclear receptor co-activator (NCOA)/steroid receptor coactivator (SRC) family and transcription-initiation-related factor KATs (Sheikh, 2014; Canovas and Ross, 2016). CREB-binding protein (CBP) and its paralog p300 were originally identified to bind the cAMP-response element binding (CREB) protein (Chrivia *et al*., 1993) and the adenovirus early-region 1A (E1A) protein (Eckner *et al*., 1994), respectively. Researchers further discovered that CBP/p300 also posess histone acetyltransferase activity

(Bannister and Kouzarides, 1996) and H3K14, H3K18, H3K27, H4K5 and H4K8 have all been identified as their targets (reviewed by Tie *et al*., 2009; Valor *et al*., 2013). CBP/p300 play important roles in cell growth and embryo development. Loss of function of CBP and p300 are both lethal in mice (reviewed by Goodman and Smolik, 2000; Philip *et al*., 2015). KATs from all five families have been found to be important in neural development (Sheikh, 2014).

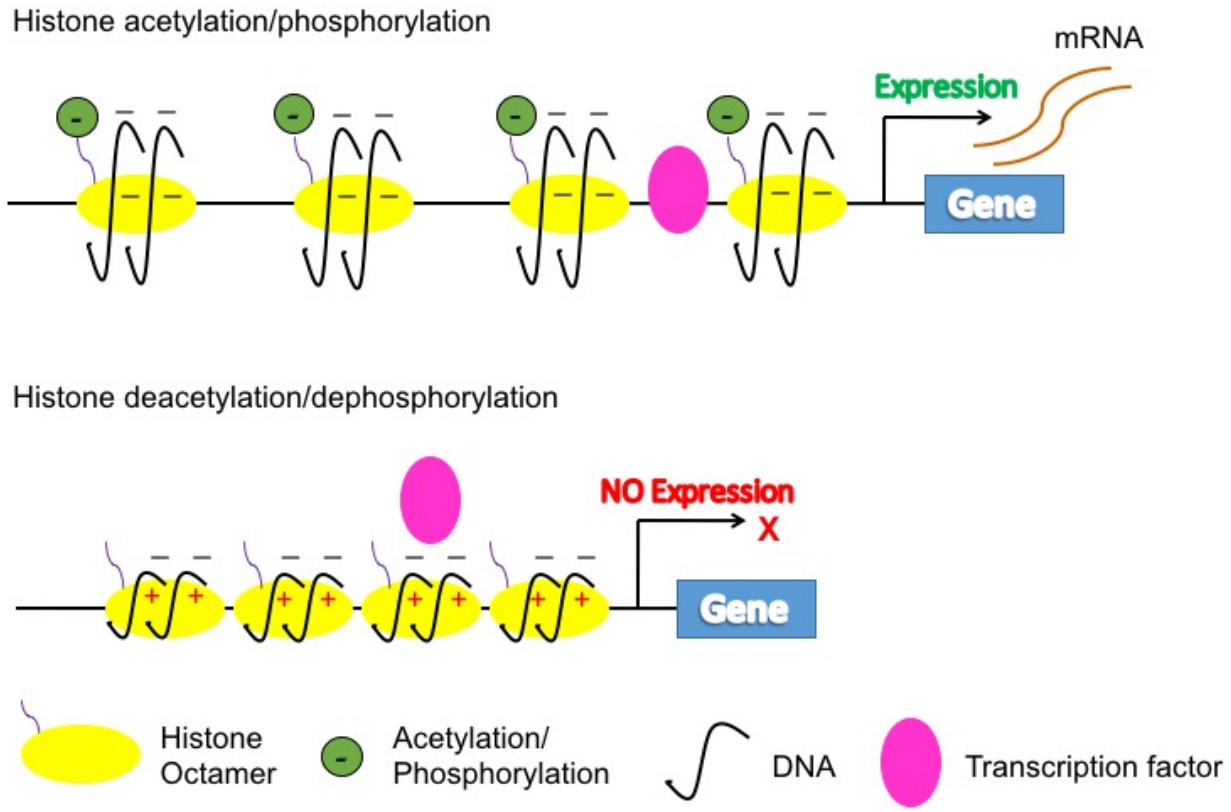


Figure 2.8. Mechanisms of gene regulation by histone acetylation and histone phosphorylation. DNA molecules are negatively charged. Histone acetylation and phosphorylation can neutralize the positive charge of histone. The interactions between histones and DNA are then weakened, making DNA accessible to transcription factors. Without acetylation/phosphorylation, chromatin is compacted and DNA is inaccessible to transcription factors.

KDACs are transcriptional repressors that reverse lysine acetylation, restore the positive charge of the lysine residue and thus stabilize the local chromatin architecture (Bannister and Kouzarides, 2011). KDACs are categorized into classical KDACs (KDAC1-11) and non-classical KDACs based on their enzymatic mechanisms. Classical KDACs all share a conserved deacetylase domain whereas the catalytic function of non-classical KDACs is NAD^+-dependent (Das Gupta *et al*., 2016). KDACs are important in embryonic development. Deletions of KDAC1, 2, or 3 are all lethal in mice (Montgomery *et al*., 2008; Guan *et al*., 2009; Dovey *et al*., 2010; Lv *et al*., 2014). KDACs are also important in cell differentiation. Inhibition of KDACs activities can lead to early differentiation of embryonic stem cells (ESCs) (Lv *et al*., 2014).

2.2.2.2 Histone phosphorylation

Similar to histone acetylation, histone phosphorylation also activates gene expression. It occurs on serines, threonines and tyrosines of the N-terminal histone tails (Rossetto *et al*., 2012; Sawicka and Seiser, 2014). Protein kinases and phosphatases are the two groups of enzymes that add and remove the modification, respectively. Histone kinases transfer a phosphate group from ATP to the hydroxyl group of the target amino-acid, add negative charge to the histone and change the chromatin structure to increase DNA accessibiliby and activate the gene (Fig. 2.8) (Bannister and Kouzarides, 2011; Brehove *et al*., 2015). Histone phosphorylation is important in DNA damage repair, transcription regulation and chromatin remodeling (Rossetto *et al*., 2012; Sawicha and Seiser, 2014).

2.2.2.3 Histone methylation

Histone methylation is associated with both transcriptional activation and gene silencing. Histone methylation occurs on lysines and arginines without changing the charge of the histone protein. There are three levels of histone lysine methylation: mono-, di- and tri-methylation (Ng *et al*., 2009; Bannister and Kouzarides, 2011). Histone lysine methyltransferase (HKMT) contain a so-called SET domain that catalyses the transfer of a methyl group from SAM to a lysine's ε-amino group (Bannister and Kouzarides, 2011; Fan *et al*., 2015). HKMTs are relatively specific enzymes with different enzymes within a family catalyzing different sites and levels of histone methylation. There are generally two classes of lysine demethylases (KDMs): the Jumonji C (JmjC) class and PHF8 (KDM7) families (Bannister and Kouzarides, 2011; Krishnan *et al*., 2011). Methylation of histone on different lysine residues can result in different regulation functions. It is reported that methylation on histone H3K4, H3K36, and H3K79 are associated with transcriptional activation while the di- and tri-methylation on histone H3K9, H3K27 induce transcriptional inhibition (Fig 2.9) (Feng *et al*., 2007; Nguyen and Zhang, 2011). Different degrees of residue methylation may result in different biological functions as well. H4K20 methylation has three states (mono-, di- and trimethylation). Mono- and di-methylation of H4K20 are involved in DNA replication in cell cycle and DNA damage repair, whie trimethylation of H4K20 is related to heterochromatin maintenance (Jørgensen *et al*., 2013).

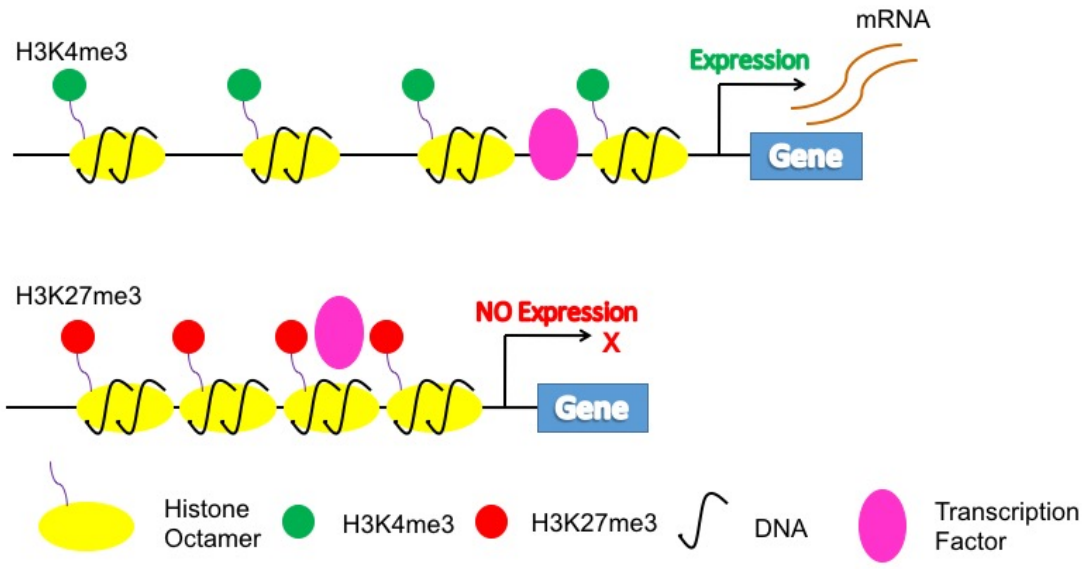


Figure 2.9. Mechanisms of gene regulation by histone methylation. Some histone methylation, like H3K4me3, can open chromatin structure and make DNA accessible to transcription factors. Other histone methylation, like H3K27me3, will result in compacted chromatin structure and DNA will not be accessible to transcription factors.

2.2.2.4 Histone modification complexes

Histone modification enzymes often form complexes with other proteins to carry their activity. These complexes are divided into two groups: the polycomb group (PcG) proteins and trithorax group (TrxG) proteins. The PcG and the TrxG proteins are epigenetic regulators responsible for the repression and activation, respectively of a group of genes important in development and cell fate specification. They are important for these genes to maintain previously established states of gene expression over multiple cell generations (Schuettengruber *et al*., 2011). The PcG and TrxG were first identified in *Drosophila* where they were shown to be required for the long-term repression and activation of the Antennapedia class of homeobox genes (Jurgens, 1985; Mallo and Alonso, 2013). The DNA regulatory elements that recruit PcG and TrxG factors to chromatin are referred to as PcG and TrxG response elements (PREs and TREs), respectively. Some TrxG and PcG components possess direct histone modification

activity while other TrxG and PcG proteins affect histone marks (Schuettengruber *et al*., 2007). PcG proteins are comprised of two types of epigenetic regulators: Polycomb Repressive Complex 1 (PRC1) and PRC2. PRC1 contains Bmi1, Ring1A, Ring1B, Cbx and Phc, and is responsible for ubiquitination of histone H2A at lysine 119, which is essential for *Hox* gene silencing (Cao *et al*., 2005; Mallo and Alonso, 2013; Piunti and Shilatifard, 2016). PRC2 is composed of Suz12, EED (embryonic ectoderm development) and Ezh1/2. Ezh2 catalyzes the di- and tri-methylation of histone H3 at lysine 27 which also acts to silence gene expression (Cao *et al*., 2002; Piunti and Shilatifard, 2016) and this mark is specifically recognized by PRC1 (Schuettengruber *et al*., 2011; Piunti and Shilatifard, 2016). There are generally two classes of TrxG: one is composed of SET domain factors and the other includes components of ATP-dependent chromatin remodeling complexes (Schuettengruber *et al*., 2011; Geisler and Paro, 2015). The SET domain histone modification complexes include COMPASS, COMPASS-like, TAC1 and absent small or homeotic discs 1 (ASH1) complexes. COMPASS and COMPASS-like complexes have H3K4 trimethylase activity and are composed of histone methytransferase (SET1A, SET1B, MLL1-4) and subunits including WD repeat domain 5 (WDR5), retinoblastoma binding protein 5 (RBBP5), absent small or homeotic 2-like (ASH2L) and DPY30. TAC1 and ASH1 complexes both coupled to CBP and counteract PcG silencing. ATP-dependent chromatin remodeling complexes includes switch/sucrose nonfermentable (SWI/SNF), imitation switch (ISWI) and various chromodomain helicase (CHD)-containing complexes that can recognize the histone tail modifications to facilitate active transcription (Schuettengruber *et al*., 2011; Geisler and Paro, 2015).

PcG and TrxG proteins play important roles in embryonic development and *Hox* gene regulation. PRC2 deficiency is fatal because single knockout of Suz12, EED or Ezh2 results in

early embryonic lethality in mice (Pasini *et al*., 2004; Faust *et al*., 1995; O'Carroll *et al*., 2001; Cheedipudi *et al*., 2014; Kadoch *et al*., 2016). Mutation of Ring1B (member of PRC1) also leads to embryonic lethality (Voncken *et al*., 2003; Morey *et al*., 2015) while other PRC1 deficiencies are less severe, only resulting in developmental abnormalities (reviewed by Jones and Wang, 2010; Morey *et al*., 2015). In ES cells, *Hox* gene promoters often have 'bivalent domains', that is, they display both H3K4me3 and H3K27me3 marks. These bivalent domains appear to keep proper expression status of *Hox* genes (Bernstein *et al*., 2005; Montavon and Duboule, 2013). PcG and TrxG protein have a "ying-yang" effect on the bivalent domains. Mutation of PcG genes results in abnormal *Hox* gene expression and leads to irregular ES cell differentiation in both human and mouse (Lee *et al*., 2006; Boyer *et al*., 2006; Morey *et al*., 2015).

2.2.2.5 Histone methyltransferase complex MLL1/WDR5

Among all the PcG and TrxG protein complexes, MLL1/WDR5 complex is of special interest in my study (discussed later in the thesis). MLL1/WDR5 complex belongs to the TrxG protein complexes and is specifically responsible for the methylation of H3K4. Mutations in MLL1 are associated with several acute lymphoblastic and myelogenous leukemias from where it derives its name. MLL1 is an H3K4 methyltransferase that belongs to the evolutionarily conserved SET1 family. In humans, this gene is located on chromosome 11 and its cDNA is ~12 kb in length. *MLL1* gene encodes a protein of ~4000 amino acids in length which can be digested in cells by taspase into two fragments: a 320 kDa N-terminal fragment and a 180 kDa C-terminal fragment (Zhang *et al*., 2013). The two cleaved peptides remain associated to form a heterodimer.

MLL1 protein contains several conserved domains which are related to its functions in chromatin-mediated transcriptional regulation. The N-terminal fragment contains three DNA-binding AT-hooks, followed by a cysteine-rich region with homology to DNA methyltransferases, referred to as the CxxC domain, followed by three plant homeodomain (PHD) zinc-finger-like motifs, and finally a bromo domain (BD) region. In the C-terminal fragment there is a transactivation domain (TAD) which can interact with CBP, a Win motif known as WDR5 binding site, and a SET domain which is responsible for the histone methyltransferase activity (Rasio *et al*., 1996; Nakamura *et al*., 2002; Zhang *et al*., 2013; Cosgrove and Patel, 2010) (Fig. 2.10A). MLL1 is a H3K4 specific methyltranferase and the CxxC domain may play a role in targeting MLL1 to active genes because this domain selectively binds to non-methylated CpG islands (Birke *et al*., 2002; Long *et al*., 2013; Bina *et al*., 2013). The PHD domains aid to recognize lysine residues on chromatins (Taverna *et al*., 2007; Lalonde *et al*., 2014). Further research has showed that the TAD domain may also contribute to the specificity of MLL1. The TAD domain of MLL1 can interact with lysine acetyltransferases CBP/p300 which contains a number of protein-binding domains that mediate transcription factor recruitment (Cosgrove and Patel, 2010; Wang *et al*., 2013). In this way, MLL1 can increase the binding of other transcriptional activators to help activate gene expression. In the TrxG, MLL1 forms complex with other components, including WDR5, RBBP5, ASH2L and DPY30 (Zhang *et al*., 2013; Shinsky *et al*., 2015). These proteins can stimulate catalytic activity and product specificity of SET1-familiar methyltransferase, including MLL1. Deletion of any of these genes in yeast leads to a similar phenotype as observed in SET1 mutants (Nagy *et al*., 2002; Ernst and Vakoc, 2012).

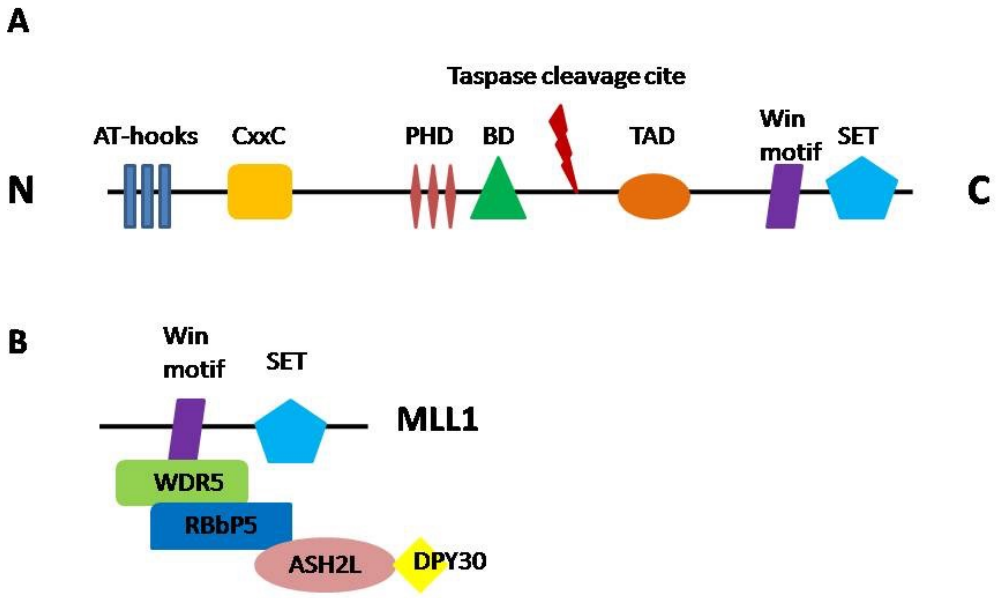


Figure 2.10. Regulation of MLL1 by the core complex subunits. (A) MLL1 domain architecture. MLL1 contains the following functional domains: AT-hooks, CxxC, PHD, BD, TAD, Win motif and SET. Taspase can cleave MLL1 between BD and TAD into two peptide fragments. (B). Interaction of important subunits in MLL1 complex (Figure modified from Ernst and Vakoc, 2012).

In mammalian cells, WDR5 seems essential for all forms of H3K4 methylation whereas knockdown of RBBP5 affects primarily H3K4me3 and H3K4me2 (Dou *et al*., 2006; Ernst and Vakoc, 2012). ASHL2 and DPY-30 have more specific role on H3K4me3 modification (Jiang *et al*., 2011; Dou *et al*., 2006; Tremblay *et al*., 2014). Research in yeast has also found that WDR5 and RBBP5 proteins are essential for the stability of SET1 as depletion of either subunit leads to SET1 degradation (Steward *et al*., 2006). WDR5 binds directly to MLL1 through the Win motif in MLL1 in a 1:1 ratio (Patel *et al*., 2008; Shinsky *et al*., 2014) and bridges MLL1 with RBBP5 and ASH2L. The hinge region in RBBP5 then binds directly to the SPRY domain of ASH2L,

and ASH2L in turn binds to DPY30 (reviewed by Ernst and Vakoc, 2012) (Fig. 2.10B). MLL1 enzyme activity is relatively weakened in the absence of any of the associated subunits mentioned above. Research has revealed that the MLL1-WDR5-RBBP5 complex has only a mild enzyme activity and is only capable of mono-methylation (Patel *et al*., 2009; Shinsky *et al*., 2014; Shinsky *et al*., 2015). Adding ASH2L to the complex leads to a significant increase in H3K4 methyltransferase activity and shifted the product specificity to di- and tri-methylation. The addition of DPY30 further increased the activity of the complex (Patel *et al*., 2009; Ernst and Vakoc, 2012).

2.2.3 MicroRNA

MicroRNAs are ~22 nucleotide (nt) RNAs first discovered in *C. elegans* that can guide a RNA silencing pathway to regulate gene expression through their effect on mRNA (Lee *et al*., 1993; Felekkis *et al*., 2010; Usmani *et al*., 2016). Currently, many studies focus on miRNA regulated gene expression. Researchers have made several advancements on the roles of miRNAs in the areas of embryo development (Yan and Jiao, 2016; Green *et al*., 2015) as well as diseases including cancer (He *et al*., 2015; Usmani *et al*., 2016), cardiovascular disease (Elia and Condorelli, 2015), neurodevelopmental disease (Ardekani and Naeini, 2010) and metabolic disease (Deiuliis 2015).

2.2.3.1 microRNA biogenesis

In vivo, miRNA is synthesized as follows (Fig. 2.11): Initially a long primary transcript of the miRNA gene (pri-miRNA) composed of at least one hairpin-like miRNA precursor is cleaved by the nuclear microprocessor complex and the endonuclease Drosha complex in the

nucleus to form 50–120 nucleotide hair-pin secondary structures, named miRNA precursors (pre-miRNA) (Ke *et al*., 2003; Kocerha *et al*., 2009; Ha and Kim, 2014). The 60- to 90-nt pre-miRNAs form the stem-loop structures which are then transported by Exportin-5 from the nucleus to the cytoplasm where the pre-miRNA hairpin stem regions are excised by the endonuclease Dicer into 18 to 22 nt double strand RNAs (Bartel, 2004; Macfarlane and Murphy, 2010; Ha and Kim, 2014). One of the two strands is the mature miRNA and the other counterpart is called miRNA*. The mature miRNAs bind to the miRNA-containing RNA-induced silencing complex (miRISC) and interact with target mRNAs mostly within the 3' untranslated region (3' UTR) based on sequence complementarity between the miRNA and its target mRNA (Maes *et al*., 2009; Ha and Kim, 2014). Argonaute (AGO), a large protein family, is a key component of RISCs (Carmell *et al*., 2002; Tang, 2005; Ha and Kim, 2014). AGO is associated with Dicer as well as the target sites of RISCs to cleave the target mRNAs (Tang, 2005; Ha and Kim, 2014). The most important determinant of miRNA function is the degree of complementarity between the proximal (5') region (also known as "seed" region or the "nucleus") of the miRNA and the mRNA (Brennecke *et al*., 2005; Felekkis *et al*., 2010; Valinezhad Orang *et al*., 2014). MiRNAs can silence gene expression by triggering the degradation of the target mRNA or blocking translation of the target mRNA (Ipsaro and Joshua-Tor, 2015; Jonas and Izaurralde, 2015). Evidances show that 66-90% of miRNA-mediated gene repressions in mammalian cells are through the degradation of target mRNAs (Eichhorn *et al*., 2014; Jonas and Izaurralde, 2015). GW182 is a key protein in the miRNA-mediated mRNA degradation. After the recognition of mRNA target by RISCs, AGO protein interacts with GW182 protein, which then binds to the cytoplasmic deadenylase complexes PAN2-PAN3 and CCR4-NOT and triggers the cellular 5′-3′ mRNA decay pathway. The deadenylated mRNAs are then decapped and finally degraded

(Valinezhad Orang *et al*., 2014; Jonas and Izaurralde, 2015). Originally researchers had proposed that miRNAs can inhibit translation at both initiation and post-initiation steps (Olena and Patton, 2010) but the use of ribosome profiling method ruled out the mechanisms occurring post-initiation (Huntzinger and Izaurralde, 2011; Jonas and Izaurralde, 2015). The molecular mechanism for the miRNA inhibited translation initiation remains to be resolved (Jonas and Izaurralde, 2015). A mRNA transcript can have multiple miRNA binding sites and be repressed simultaneously by different miRNA's. The same or different miRNAs bind within the same 3′UTR and can act cooperatively to enhance repression (Doench and Sharp, 2004; Felekkis *et al*., 2010).

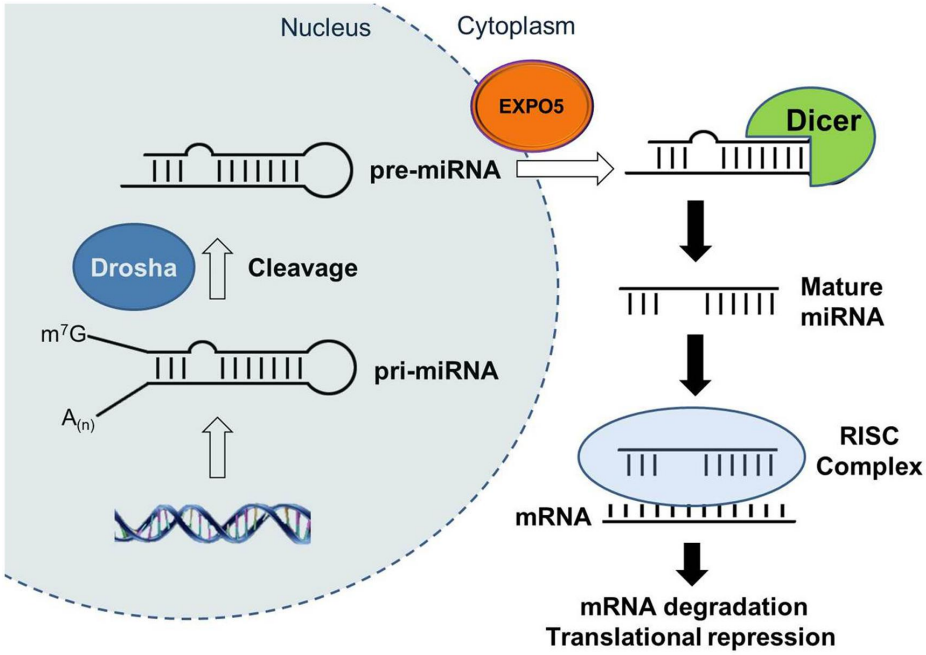


Figure 2.11. miRNA biogenesis and function. Transcribed by RNA polymerase II, long primary transcripts (pri-miRNA) are then processed in the nucleus into hairpins (pre-miRNA) by Drosha and DGCR8. Exportin 5 exports the pre-miRNA to cytoplasm where it is cleaved into a mature dsRNA duplex. One strand that is stably associated with Argonaute (AGO) is packaged into the RISC to target and regulate mRNAs. The other strand is unwound and degraded. The target mRNA is either degraded or translation is repressed. This figure is taken with permission from from Jung and Suh (2015).

Table 2.3. Regulation of *Hox* genes by miRNAs

Species	miRNA	*Hox* Gene Target	System impacted	Reference
Drosophila	miR-iab-4	Ubx, Abd-A	Halteres to wing transformation	Ronshaugen *et al*., 2005
	miR-iab-8	Ubx, Abd-A	Halteres to wing transformation	Tyler *et al*., 2008
		Ubx	CNS	Thomsen *et al*., 2010
		Abd-A	CNS	Gummalla *et al*., 2012
		Ubx, Abd-A	CNS, Fertility	Garaulet *et al*., 2014
Zebrafish	miR-10	Hoxb1a, Hoxb3a	Hindbrain	Woltering and Durston, 2008
	miR-196	Hoxb8a	Body axis development	He *et al*., 2011b
Chick	miR-196	Hoxb8	Hindlimb development	Brock *et al*., 2009
			Skeletogenesis	McGlinn *et al*., 2009
	miR-130a	Hoxa3	Tumorigenesis	Han *et al*., 2016
Mouse	miR-196	Hoxb8	Hindlimb development	Brock *et al*., 2009
		Hoxa7,Hoxc8, Hoxb8	Vertebral number and vertebral identity	Wong *et al*., 2015
	miR-181	Hoxa11	Skeletal muscle cell differentiation	Naguibneva *et al*., 2006
	miR-3960	Hoxa2	Primary osteoblasts	Hu *et al*., 2011
	miR-130a	Hoxa5	Airway branching and lung microvascular development	Mujahid *et al*., 2013
	miR-221	Hoxb5		
Human	miR-196	HOXB8, HOXC8, HOXD8, HOXA7	Cervical cancer (Hela) cells	Yekta *et al*., 2004
	miR-196a	HOXB9	Head and neck squamous cell carcinoma	Darda *et al*., 2015
	miR-196b	HOXA9	Lung cancer	Yu *et al*., 2015
			Chronic myeloid leukemia	Liu *et al*., 2013
			MLL-rearranged leukemia	Li *et al*., 2012
	miR-126, miR-145, let-7	HOXA9	Bone marrow	Shen *et al*., 2008
	miR-193a-3p	HOXC9	Bladder cancer	Lv *et al*., 2015
	miR-135a	HOXA10	Epithelial ovarian cancer	Tang *et al*., 2014
	miR-100	HOXA1	Small cell lung cancer	Xiao *et al*., 2014
	miR-99	HOXA1	Epithelial cell proliferation and migration	Chen *et al*., 2013
	miR-10	HOXD10	Non-metastatic breast cancer	Ma *et al*., 2007
			Ovarian cancer	Nakayama *et al*., 2013
	miR-130a	HOXA5	Human umbilical vein endothelial cells	Chen and Gorski, 2008
	miR-221	HOXB5	Papillary thyroid carcinoma	Kim *et al*., 2008
	miR-7, miR-218	HOXB3	Breast cancer	Li *et al*., 2012

2.2.3.2 Regulation of Hox genes by miRNAs

Studies show that miRNAs can regulate the expression of *Hox* genes and play important roles in embryonic development. Some miRNA genes are transcribed from *Hox* gene clusters like mir-10, mir-196 and iab-4 (Lagos-Quintana *et al*., 2011; Aravin *et al*., 2003; Woltering and Durston, 2008; Yekta *et al*., 2004; Hornstein *et al*., 2005). These miRNAs are expressed in patterns that approximate the characteristic expression of *Hox* genes. Other miRNAs have also been found to regulate *Hox* genes. Below I have summarized research on miRNAs and *Hox* gene regulation that has so far been investigated in several species including insects, mouse, chicken, fish and human cell lines (see Table 2.3).

In *Drosophila,* three homeotic genes within BX-C encode homeobox proteins. They are Ultrabithorax (Ubx), Abdominal-A (Abd-A) and Abdominal- B (Abd-B) (Martin *et al*., 1995; Garaulet and Lai, 2015). Twe groups of *Drosophila* miRNAs are transcribed from BX-C and can regulate Hox gene expression (Garaulet and Lai, 2015). The iab-4–5p and iab-4–3p are located at the end of the iab-4 locus between *abd-A* and *Abd-B* genes (Aravin *et al*., 2003; Garaulet and Lai, 2015). Ronshaugen *et al.*, (2005), demonstrated direct targeting of miR–iab-4–5p in Ubx 3′ UTR and showed that iab-4 miRNA down regulates Ubx activity *in vivo.* Over expression of mir-iab-4-5p attenuates Ubx protein levels and induces the transformation of halteres into wings (Ronshaugen *et al*., 2005). iab-4 miRNA was also predicted to repress the downstream targets of the homeobox genes (Grun *et al*., 2005). Thus iab-4 miRNA may fine-tune the development of *Drosophila* via both homeobox genes and their downstream targets (Chopra and Mishra, 2006). The miRNA iab-8 is a transcript produced on the sense strand of the coding region of BX-C but antisense to iab-4 (Tyler *et al*., 2008; Gummalla *et al*., 2012; Garaulet and Lai, 2015). Ubx appears to be a much stronger target of miR-iab-8-5p than of miR-iab-4-5p (Tyler *et al*., 2008;

Garaulet and Lai, 2015). Ectopic expression of iab-8 leads to haltere to wing transformation which is similar to what occurs after over expression of mir-iab-4-5p (Tyler *et al*., 2008; Bender, 2008). *Abd-A* is also regulated by iab-8 where deletion of mir-iab-8 influenced the expression of both *Ubx* and *Abd-A* in the posterior larval ventral nerve cord (VNC) and disrupted CNS patterning and reproductive behavior in *Drosophila* (Garaulet *et al*., 2014).

miRNAs are also very important in *Zebrafish* development. It has been shown that Dicer1, the microRNA-producing enzyme, is essential for *Zebrafish* development. *Zebrafish* embryos with blocked Dicer activity display abnormal morphogenesis during gastrulation, brain formation, somitogenesis, and heart development (Giraldez *et al*., 2005; He *et al*., 2011a). miR-10 gene is located 5′ of the Hox-4 genes and binds to *Zebrafish Hox* genes. In *Zebrafish*, miR-10c is located on the same primary transcript as *Hoxb4a* and is expressed in a Hox-4 like pattern in the spinal cord (Woltering and Durston, 2008; Mallo and Alonso, 2013). Binding sites of miR-10 have been experimentally confirmed to occur in both *Zebrafish Hoxb1a* and *Hoxb3a* 3′ UTR and in the open reading frame (Woltering and Durston, 2008). It has been reported that miR-10 represses *Hoxb1a* and *Hoxb3a* gene expression and its overexpression leads to loss of function phenotypes for both genes (Woltering and Durston, 2008). miR-196 is another *hox-cluster* miRNA that down regulates *hoxb8a* expression in *zebrafish* by targeting its 3′ UTR. An over expression or knockdown of miR-196 leads to abnormal embryo development in *zebrafish* (He *et al*., 2011b; Mallo and Alonso, 2013).

In cultured human Hela cells, miR-196 down regulates the expression of *HOXB8, HOXC8, HOXD8,* and *HOXA7* by binding to 3′ UTR regions (Yekta *et al*., 2004). In patients with chronic myeloid leukemia, lower levels of miR-196b expression were detected and further experiments showed that miR-196b targets *HOXA9* and reduces cell proliferation rate (Liu *et al*.,

2013). miR-196b mediated regulation of *HOXA9* expression has also been reported in mesenchymal-like-state non-small cell lung cancer cells (Yu *et al*., 2015). HOXB9 is also regulated by miR-196a in head and neck squamous cell carcinoma (Darda *et al*., 2015). Other experiments showed targeting of *HOXD10* mRNA by miR-10 in cultured human non-metastatic breast cancer (SUM149) cells (Ma *et al*., 2007). Luciferase assays in HCT 116 cells demonstrated that miR-10a strongly repressed luciferase activities mediated by 3′ UTR vectors derived from *HOXA3* and *HOXD10* (Han *et al*., 2007). In ovarian cancer cells, over expression of miR-10b down regulated the expression of *HOXD10* and accelerated the migration and invasion activities of these cancer cells (Nakayama *et al*., 2013). HOXA10 is another gene important in epithelial ovarian cancer (EOC). A down regulation of HOXA10 by miR-135a can lead to enhanced cell apoptosis and inhibition of cell growth and adhesion in EOC-derived cell lines (Tang *et al*., 2014). Down regulation of both antiangiogenic homeobox genes *GAX* and *HOXA5* by miR-130a was observed in human umbilical vein endothelial cells (HUVECs), though the antagonism is weaker for *HOXA5* than what was observed for *GAX* (Chen and Gorski, 2008). miR-221 is highly expressed in human papillary thyroid carcinoma (PTC) cell lines and altered the expressions of many target genes (Kim *et al*., 2008). HOXB5 is one of the most significantly affected gene and was identified as a direct target of miR-221 (Kim *et al*., 2008). miR-126, miR-145, and let-7 are all found to target the full-length *HOXA9* cDNA and can block *HOXA9* biological functions. More interestingly, miR-126 regulates *HOXA9* by binding to the homeobox region rather than the 3′ UTR (Shen *et al*., 2008). Two tumor suppressor microRNAs, miR-7 and miR-218, were found to target *HOXB3* in human breast cancer cell lines and further regulate the expression of tumor suppressor genes RASSF1A and Claudin-6 (Li *et al*., 2012). Three members of miR-99 family (miR-99a, miR-99b and miR-100) can bind directly

to *HOXA1* mRNA. An over expression of miR-99 family led to reduced *HOXA1* expression and caused reduced proliferation and cell migration (Chen *et al*., 2013; Xiao *et al*., 2014). miR-193a-3p can directly regulate the expression of HOXC9 and plays an important role in bladder cancer chemoresistance. In bladder cancer cell lines, ectopic expression of miR-193a-3p reduced HOXC9 expression and led to chemoresistance to several drugs (Lv *et al*., 2015).

In mouse embryos, *Hoxb8* is directly cleaved by miR-196 (Yekta *et al*., 2004). miR-196 was also shown to regulate *Hoxb8* in both mouse and chick hindlimb development (Brock *et al*., 2009). Observations by McGlinn and colleagues showed that loss of function of miR-196 in chick embryos led to extensive skeletal defects including homeotic transformations (McGlinn *et al*., 2009). *Hox* genes, *Hoxa7* and *Hoxc8*, have also been identified as targets of miR-196 in mouse embryo development (Wong *et al*., 2015). miR-130a regulates Hoxa3 expression at the protein level in the chick and is involved in the regulation of cell proliferation in tumorigenesis associated with the chicken Marek's disease (Han *et al*., 2016). miR-181 is strongly upregulated during differentiation of skeletal muscle cells. miR-181 can downregulate *Hoxa11* protein expression and is suggested to play a role in skeletal-muscle differentiation (Naguibneva *et al*., 2006). In mouse embryo development, miR-130a and miR-221 were shown to regulate both airway branching and lung microvascular development by targeting *Hoxa5* and *Hoxb5*, respectively (Mujahid *et al*., 2013). While a significant number of miRNAs have been reported to regulate *Hox* gene expression, only one miRNA has been reported to directly target *Hoxa2*. Hu and colleagues identified a novel miR-3960 from primary mouse osteoblasts and instead of targeting the 3′ UTR region, it targets *Hoxa2* coding region. Together with miR-2861 which targets the *Hoxa2* downstream gene, *Runx2*, miR-3960 can form a regulatory feedback loop and

plays an important role in osteoblast differentiation (Hu *et al*., 2011). Whether there are any specific miRNAs that target Hoxa2 3′ UTR still remain to be discovered.

2.2.4 Long noncoding RNA (lncRNA)

Although 70-90% of the mammalian genome is believed to be transcribed, only 1% of the genome encodes protein (Lee, 2012). Researchers previously considered these noncoding transcripts as transcriptional noise, but are beginning to realize that many of these noncoding transcripts have functions, and are now classified as long noncoding RNAs (lncRNAs). LncRNAs are a group of RNAs that do not encode proteins, and are comparably longer than other short noncoding RNAs, such as microRNAs and tRNAs. A majority of lncRNAs share the same transcriptional machinery with mRNAs, namely; (1) their transcription is catalyzed by RNA polymerase II; (2) lncRNAs also have 5' methylguanosine cap and 3' polyadenylation and often undergo splicing (Mercer and Mattick, 2013). However, lncRNAs usually lack an extended open reading frame (ORF). It is difficult to determine the exact number of human lncRNAs. Researchers estimate there are at least 5,000 to 15,000 long noncoding transcripts with many lncRNAs yet to be annotated (reviewed by Mercer and Mattick, 2013). Although there are a large number of lncRNAs, individual lncRNA tend to express in significantly lower quantities compared to their protein-coding counterparts, making these difficult to detect (Cabili *et al*., 2011). LncRNAs can be transcribed between or within coding genes. Many lncRNAs are transcribed from the antisense of coding genes (Mercer and Mattick, 2013).

LncRNAs have emerged with important regulatory roles in gene expression. An important question that needs addressing is, how do lncRNAs regulate gene expression? At the structural level, lncRNAs carry their regulatory function by interacting with other molecules,

including DNA, RNA and protein. Proteins are the most common partner of lncRNAs, with at least 15% of expressed proteins having RNA binding capacity (Baltz *et al*., 2012). Proteins usually do not recognize a specific lncRNA sequence but tend to interact with their secondary structures (Mercer and Mattick, 2013). Significant evidence indicates that lncRNAs can recruit chromatin-modifying proteins to specific promoter sites. The indication that lncRNAs may be crucial accessory factors for Polycomb function first arose when researchers found the interacting factors of Xist: PRC2 (Zhao *et al*., 2008). In fact, many histone methyltransferases (HMTases) lack DNA binding domain but usually have RNA binding capacity, indicating lncRNAs may play an important role in guiding HMTases to certain chromatin loci (Bernstein and Allis, 2005). Some lncRNAs can also interact with mRNAs. For example, AS Uchl1 is a lncRNA anti-sense to mouse ubiquitin carboxyterminal hydrolase L1 (*Uchl1*). It can interact with *Uchl1* mRNA to regulate its stability and translation (Carrieri *et al*., 2012). Although there is currently little evidence for direct binding of lncRNA to DNA, researchers have reported a lncRNA, the pRNA (promoter-associated RNA) which forms a DNA:RNA triplex with ribosomal DNA (rDNA) promoter to be specifically recognized by the DNA methyltransferase DNMT3b (Schmitz *et al*., 2010). LncRNAs are biologically important in many processes. Below I will review several examples of lncRNAs that have important biological functions.

2.2.4.1 LncRNAs and X-Inactivation

In mammals to balance the dosage of X-linked genes between the two genders, a majority of genes on one X-chromosome are inactivated in females, known as X-chromosome inactivation (XCI) (Lyon, 1961; Yue *et al*., 2015). This inactivation process is driven by a series of lncRNAs, including *Xist*, *Tsix*, *RepA* and *Jpx* (Froberg *et al*., 2013; Yue *et al*., 2015). *Xist* is

one of the first lncRNAs identified to induce chromosome-wide silencing on inactive X (Xi) (Penny *et al*., 1996). In the active X (Xa), *Tsix* is expressed which is antisense to *Xist* and can block the expression of *Xist* (Lee, 2000). As mentioned earlier, *Xist* can silence X-chromosome by recruiting PRC2 to specific gene sites. At the 5′ end of *Xist*, there is a repeat motif identified as "Repeat A" (Zhao *et al*., 2008). This motif directly interacts with EZH2, the histone methyltransferase in PRC2, which trimethylates histone H3 at lysine 27. With the help of *Xist*, PRC2 is initially localized to "PRC2 strong sites" (e.g. the bivalent domains and CpG islands) along the Xi. It then migrates to other parts of the chromosome and eventually silences the entire chromosome (Pinter *et al*., 2012; Valencia and Wutz 2015; Goodrich *et al*., 2016). A recent study also showed that *Xist* can silence transcription through KDAC3. Moreover, KDAC3 is critical for *Xist*-mediated recruitment of PRC2 across Xi (McHugh *et al*., 2015). Expression of *Xist* is regulated by other noncoding RNAs. *Jpx* and *Ftx* are two noncoding RNAs encoded upstream of Xist promoter that act as activators of Xist (Nora *et al*., 2012; Goodrich *et al*., 2016). In Xa, *Tsix* controls *Xist* expression by methylating its promoter and modulating its chromatin structure (Sado *et al*., 2005). Evidence showed that *Tsix* can interact with PRC2 and establishes H3K27me3 on *Xist* as well as recruit DNMT3a to enhance hypermethylation on *Xist* promoter (Sado *et al*., 2005; Ohhata *et al*., 2008; Furlan and Rougeulle, 2016).

2.2.4.2 LncRNAs from *Hox* loci

There are several lncRNAs that have been identified to be transcribed from *Hox* clusters (Fig. 2.12). Investigation on the transcriptional activity of the human *HOX* loci showed that many of the intergenic regions are actively transcribed and most of the transcripts from the intergenic regions are lncRNAs (Rinn *et al*., 2007; De Kumar and Krumlauf, 2016). Up to 74%

of these lncRNAs are transcribed from the opposite-strand of *HOX* genes. Rinn and colleges concluded that *Hox* lncRNAs play important biological roles based on the following key facts: 1. Some lncRNAs are conserved in evolution; 2. Like *HOX* genes, these lncRNA have different expression patterns along the A-P axes depending on their physical location on the chromosome; 3. These lncRNAs possess specific sequence motifs based on their site-specific expression patterns (Rinn *et al*., 2007; De Kumar and Krumlauf, 2016).

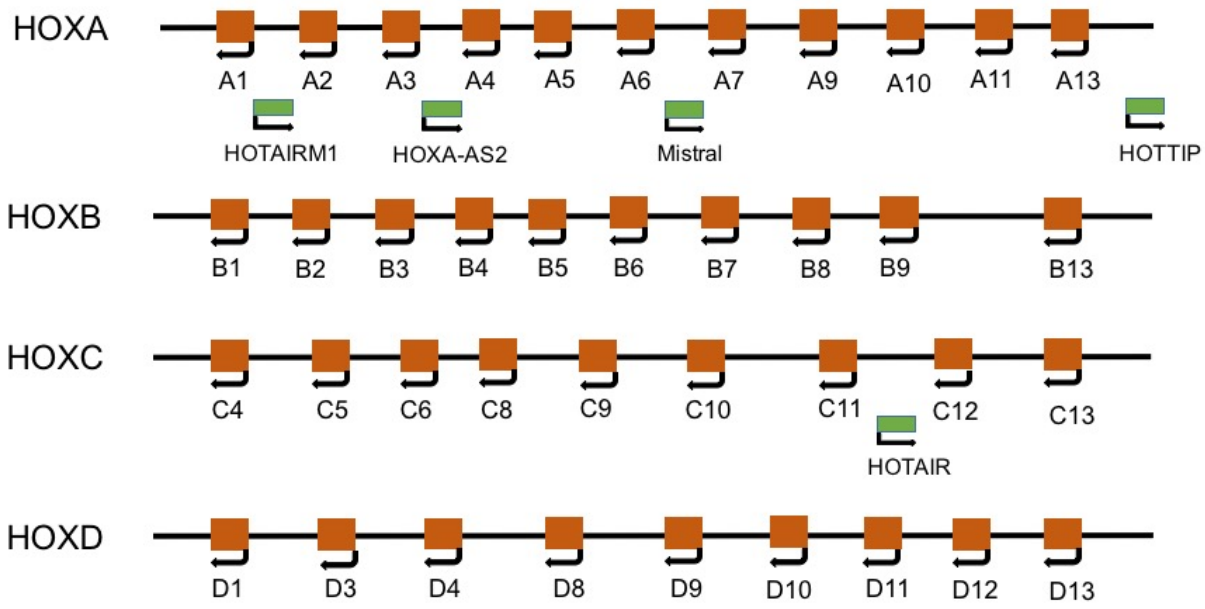


Figure 2.12. Noncoding transcripts originating from the four Hox clusters. Relative positions of non-coding transcripts (green) are shown on the basis of human Hox coding genes (red) as landmark. Arrows indicate the direction of transcription.

Table 2.4. Summary of lncRNAs transcribed from the HOX gene cluster.

LncRNA	Binding partner	Epigenetic marks induced	Target gene regulation		Functions and Systems	References
HOTAIR	PRC2 (Suz12 and EZH2)	H3K27me3	HOXD8, HOXD9, HOXD10, HOXD11	↓	Human primary foreskin fibroblast cells	Rinn *et al*., 2007
	—	DNA methylation	PTEN	↓	Human laryngeal squamous cell cancer	Li *et al*., 2013
	PRC2 (EZH2)	H3K27me3	P21	↓	Cell cycle disorder	Liu *et al*., 2016a
	PRC2	H3K27me3	WIF-1	↓	Activate Wnt pathway in cancer cells	Ge *et al*., 2013; Li *et al*.,2016
HOTTIP	TrxG (WDR5)	H3K4me3	Hoxa10,Hoxa11, Hoxa13	↑	Mouse and chick embryo development	Wang *et al*., 2011
	—	—	P21	↓	Regulate cell growth and apoptosis in colorectal cancer	Lian *et al*., 2016
	—	—	cyclin D1, CDK4 and β-catenin	↑	Activate the Wnt/β-catenin pathway in osteosarcoma	Li *et al*., 2016
Mistral	TrxG (MLL1)	H3K4me3	Hoxa6, Hoxa7	↑	Stem cell differentiation	Bertani *et al*., 2011
HOTAIR M1	—	—	HOXA1, HOXA4, HOXA5	↑	Myeloid differentiation and maturation	Zhang *et al*., 2009; Zhang *et al*., 2014
HOXA-AS2	PRC2	H3K27me3	P21, PLK3, DDIT3	↓	Gastric cancer	Xie *et al*., 2015

HOTAIR was the first lncRNA to be studied in human *HOX* clusters. *HOTAIR* is a HOX-C loci antisense transcript expressed in posterior and distal sites (Rinn *et al*., 2007). It has been reported to bind to Suz12 and EZH2, components of PRC2 and induces H3K27me3 in HOXD locus in human primary foreskin fibroblast cells (Rinn *et al*., 2007). *HOTAIR* expression is related to several types of cancers. *HOTAIR* promoter contains estrogen response elements and its expression can be induced by estradiol in the human breast cancer cell line MCF7 (Bhan *et al*., 2013). *HOTAIR* is essential for the growth of many cancer cells and knockdown of *HOTAIR* induces apoptosis in these cancer cells (Bhan *et al*., 2013; Kong *et al*., 2015; Tang *et al*., 2016). In human laryngeal squamous cell carcinoma cell line Hep-2, HOTAIR suppresses the expression of *PTEN*, a tumor suppressor gene, via DNA methylation (Li *et al*., 2013). Knockdown of *HOTAIR* in Hep-2 cells also leads to apoptosis and can inhibit the invasion of Hep-2 cells (Li *et al*., 2013). Other genes and pathways regulated by *HOTAIR* in cancer cells include p21, NF-κB and Wnt/β-catenin pathway (Liu *et al*., 2016a; Özeş *et al*., 2016; Li *et al*., 2016). Expression of *HOTAIR* also increased proliferation and invasion of several tumor cells (Fang *et al*., 2016, Lee *et al*., 2016). *HOTAIR* is also involved during heart failure (Greco *et al*., 2016).

HOTTIP is an antisense transcript transcribed from 5' to *HOXA13* and is required for the activation of 5' HOXA genes (Wang *et al*., 2011). *HOTTIP* is expressed in distal human fibroblasts and in posterior sites of mouse and chick embryos. Knockdown of *HOTTIP* in chick embryo causes shortened forelimbs. ChIP-seq results showed that the down regulation of *HOTTIP* decreased H3K4me3 occupancy in 5' HOXA genes, from where *HOTTIP* is transcribed. Further results showed that *HOTTIP* can bind directly to WDR5, a protein interacting with histone methyltransferase MLL in TrxG. *HOTTIP* is essential for the recruitment

of MLL1 to 5' HOXA genes through WDR5 (Wang *et al*., 2011). Like *HOTAIR*, *HOTTIP* also plays a very important role in cancer but through regulation of different HOX genes. *HOTTIP* regulated *HOXA13* expression contributes to the progression of several cancers (Ren *et al*., 2015; Li *et al*., 2015b; Zhang *et al*., 2016). *HOTTIP* can also regulate the expression of p21 and Wnt/β-catenin pathway and enhance tumor cell proliferation and migration (Li *et al*., 2016; Cheng *et al*., 2015; Lian *et al*., 2016).

Mistral is another lncRNA found in HoxA cluster. It is an antisense transcript located between *Hoxa6* and *Hoxa7* in mice (Bertani *et al*., 2011). Similar to HOTTIP, it can activate nearby gene expression (*Hoxa6* and *Hoxa7*) by interacting with MLL1. In this way it induces H3K4me3 modification in *Hoxa6* and *Hoxa7* and this process is important in stem cell differentiation (Bertani *et al*., 2011).

A model has been suggested by Rinn *et al*. (2007) that transcription of lncRNA in *cis* (e.g. *HOTTIP*) may recruit TrxG proteins such as MLL1 and WDR5 to chromatin, leading to H3K4me3 and gene activation (Fig. 2.13A). While lncRNAs transcribed but acting in *trans* may interact with PRC2 and result in H3K27me3 modification and gene silencing (Fig. 2.13B). The later findings on *HOTTIP* and *Mistral* support this model.

Other lncRNAs identified from the *Hox* cluster include *HOTAIRM1* and *HOXA-AS2*. *HOTAIRM1* is an antisense intergenic transcript transcribed between human *HOXA1* and *HOXA2* (Zhang *et al*., 2009). It can activate 3' HOXA genes and its presence is necessary during myeloid differentiation and myeloid maturation (Zhang *et al*., 2009; Zhang *et al*., 2014). *HOXA-AS2* is a lncRNA located between the *HOXA3* and *HOXA4* genes in the HOXA cluster (Zhao *et al*., 2013). Its expression was found in NB4 promyelocytic leukemia cells and human peripheral blood neutrophils. *HOXA-AS2* expression is responsive to all-*trans* retinoic acid (ATRA)

induction in NB4 cells and can negatively regulate ATRA-induced NB4 cell apoptosis by regulating caspase 8 and 9 pathway (Zhao *et al*., 2013). *HOXA-AS2* can also recruit PRC2 to P21, PLK3 and DDIT3 genes to induce H3K27me3 and silence their expression in gastric cancer cells (Xie *et al*., 2015).

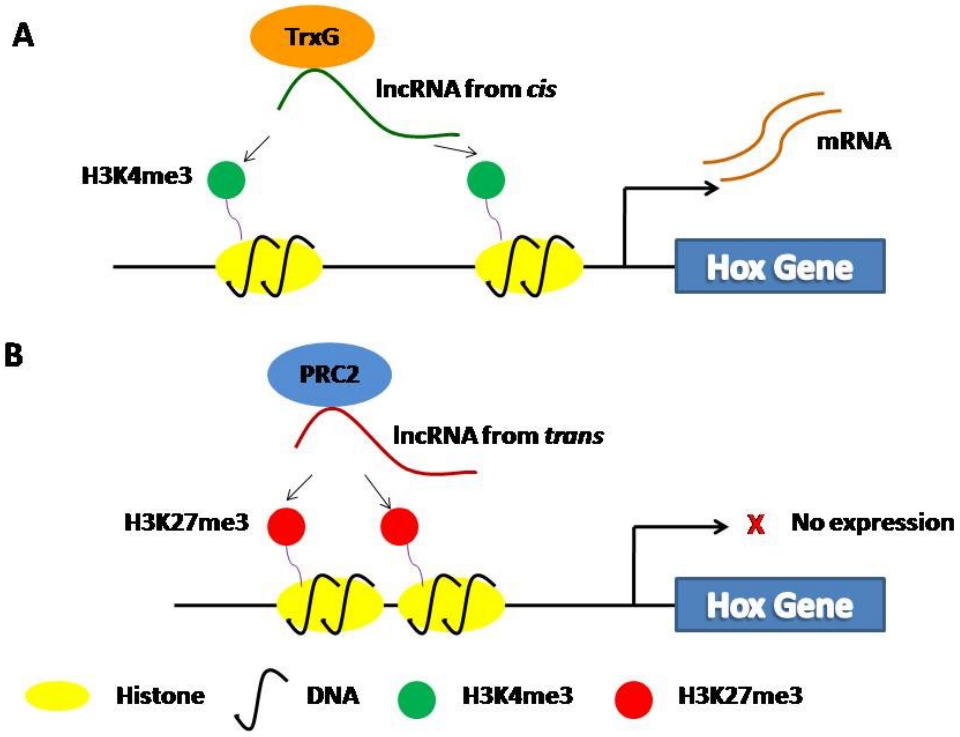


Figure 2.13. Model of long ncRNA regulation of chromatin domains via histone-modification enzymes. (A) Transcription of lncRNAs in *cis* may recruit TrxG proteins to chromatin, leading to H3K4 tri-methylation, open the chromatin structure and transcriptionally activate neighboring HOX genes. (B) lncRNAs acting in *trans*, in contrast, can recruit PRC2 to chromatin, causing a compacted chromatin structure and lead to H3K27 methylation and gene repression.

2.2.5 Cross talk between epigenetic regulators

Thus far I have reviewed how each individual epigenetic regulator can influence gene expression. The reality is usually that the regulation of gene expression is not controlled simply by one epigenetic regulator but is interrelated with many different epigenetic regulators having a coordinated affect on gene expression. Hence, more often multiple epigenetic regulators are involved in regulating the expression of a single gene. One good example is the regulation of gene expression via lncRNA mediated histone methylation that was reviewed above. Histone methylation and histone acetylation are also closely related. It has been reported that TrxG proteins can interact directly with KAT CBP and monomethylation of H3K4 can enhance H3K27 acetylation through CBP (Tie *et al*., 2014).

In addition, epigenetic regulators have the ability to affect each other. First, the ability of DNA methylation and histone modification to regulate gene expression suggests that the expression of lncRNA and miRNA may be regulated by these modifications (Han *et al*., 2007). It is predicted that over 90% of the human miRNA promoters are located 1,000 bp upstream of the mature miRNA (Zhou *et al*., 2007). In human colorectal cancer cell line CRC, hsa-miR-9 gene is included within 1,000 bp of a CpG island which is hypermethylated (Bandres *et al*., 2009). The hsa-miR-9 expression is downregulated in primary CRC compared to matched normal colorectal epithelial tissues (Bandres *et al*., 2006). Both DNA demethylation and H3 deacetylation upregulate the expression of the mature hsa-miR-9 (Bandres *et al*., 2009). Han *et.al.* (2007) have reported 13 miRNAs in colon cancer cells HCT116 that are regulated by DNA methylation. Moreover, a high level of CpG site demethylation is essential for the re-expression of these miRNAs (Han *et al*., 2007). When comparing patients with leukemia to healthy individuals, a higher degree of DNA methylation in the CpG islands of miR-196b promoter and a lower

expression of miR-196b was observed in patients with chronic myeloid leukemia (Liu *et al*., 2013). Research has shown that miR-127 is induced from its own promoter after treatment with demethylating drug 5-aza-20-deoxycytidine (5-Aza-CdR) and 4-phenylbutyric acid (PBA) in both LD419 and T24 cells. This induction is observed together with a decrease in DNA methylation levels and an increase in histone H3 acetylation as well as histone H3K4 methylation (Saito *et al*., 2006). Scott and colleagues observed that when they inhibited histone deacetylase, 40% of tested miRNAs showed significant changes in expression level within 5 h of treatment (Scott *et al*., 2006). LncRNA *HOTAIR* mediates changes in histone methylation status on miR-205 promoter to regulate miR-205 expression in bladder cancer cells (Sun *et al*., 2015).

While DNA methylation and histone modification can regulate the expression of miRNAs and lncRNAs, several researchers have found that miRNAs can also influence DNA methylation and histone modification. First, miRNAs themselves can target genes that control epigenetic pathways. Clues have been found for miR-140 to target histone deacetylase 4 in mice (Tuddenham *et al*., 2006). Ezh2, described earlier as a component of PRC2, has its expression regulated by miR-26a, miR-101, miR-205 and miR-214 (reviewed by Sato *et al*., 2011). Bmi1 belongs to PRC1 and is regulated by miR-128 (Godlewski *et al*., 2008) and mir-203 (Wellner *et al*., 2009). In prostate cancer, down regulation of miR-449a causes overexpression of KDAC-1 (Noonan *et al*., 2009). miR-29 family of miRNAs are shown to directly target DNMT 3a and 3b (Fabbri *et al*., 2007). miR-142-3p downregulates the expression of TrxG proteins ASH1L and MLL1 in thyroid follicular tumorigenesis (Colamaio *et al*., 2015). An intriguing experiment presented that miR-165 and miR-166 can interact with the newly processed PHABULOSA (PHB) mRNA and induce DNA methylation (Bao *et al*., 2004). Another research study showed that the accumulation of miRNA: target-RNA duplexes hypermethylate the genes encoding

target RNAs, leading to gene silencing (Khraiwesh *et al*., 2010). miRNAs can also regulate lncRNA expression. For example, iab-4 non-coding RNA is a substrate for miRNA (Bender, 2008). miR-192 and miR-204 can directly suppress expression of HOTTIP in hepatocellular carcinoma (Ge *et al*., 2015).

2.3 WDR5

The WD40 repeat (also known as the WD or beta-transducin repeat) proteins are a group of proteins containing a highly conserved core/repeating units of approximately 40 amino acids that usually terminate with tryptophan (Trp, W) - aspartic acid (Asp, D) dipeptide (Neer *et al*., 1994). WD repeat proteins normally contain 4-16 WD repeats (Li and Roberts, 2001). The WD repeat proteins are a large family of proteins present in all eukaryotes and involved in a variety of cellular activities including cell division (Li *et al*., 2015a), transmembrane signalling (Neer *et al*., 1994), autophagy (Grimmel *et al*., 2015), gene regulation (Sun *et al*., 2015) and apoptosis (Li *et al*., 2015a). These proteins play key roles in the formation of protein-protein complexes in nearly all the major pathways and organelles unique to eukaryotic cells, associating them with many genetic diseases (Smith, 2008).

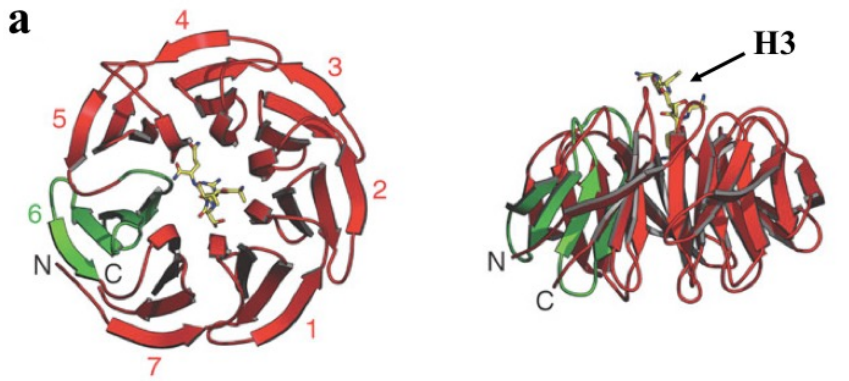
WDR5 is a protein belonging to the WD repeat protein family. It was first identified in 2001 and named BIG-3 (BMP-2 induced gene 3kb, Gori *et al*., 2001). Later it was renamed WDR5 (WD repeat protein 5) as it contains seven WD-40 repeats (Gori *et al*., 2005). As a member of the WD repeat protein family, WDR5 is hypothesized to have the three common features shared by WD repeat proteins (Li and Roberts, 2001) and these are: 1) the WD repeat domains are folded into beta propellers; 2) the WD repeat domains can reversibly assemble to multiple protein complexes and serve as a platform without any catalytic activities; 3) WDR5

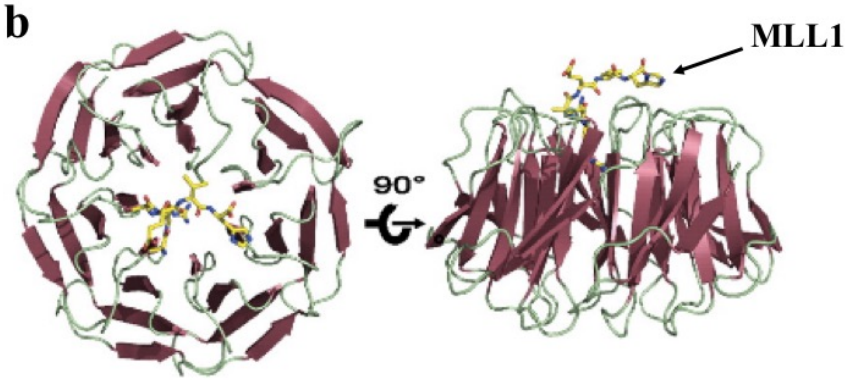
may be involved in regulating cellular functions, such as cell division, cell-fate determination, gene transcription, and mRNA modification. In fact, WDR5 is linked to various diseases including leukemia (Ali *et al*., 2014), breast carcinogenesis (Dai *et al*., 2015), prostate cancer (Kim *et al*., 2014) and pulmonary hypertension (Chen *et al*., 2015).

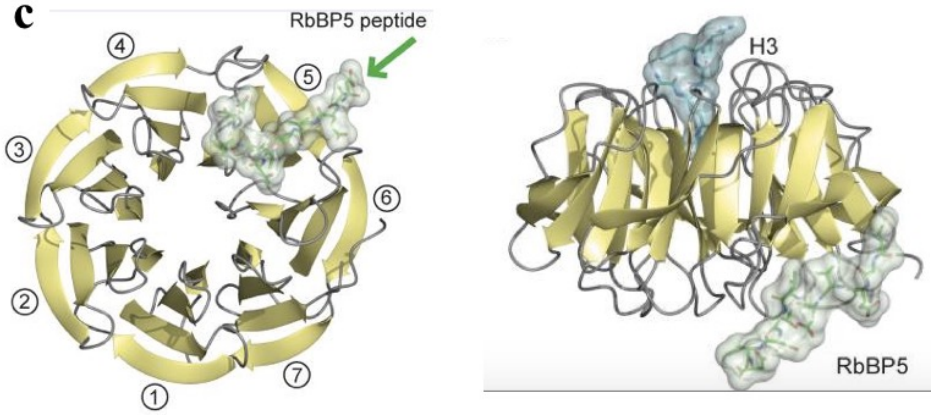
2.3.1 Structure of WDR5

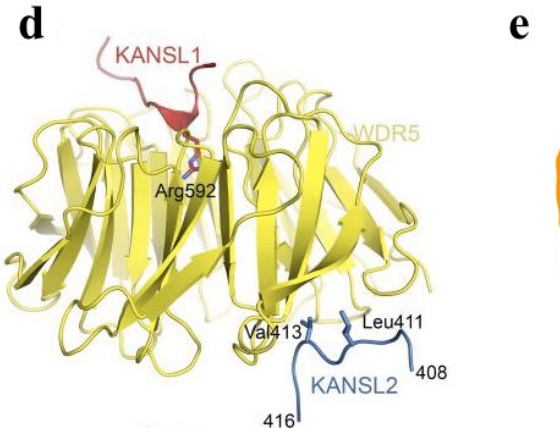
Since the first structural determination of WD repeat protein in 1995 (Wall *et al*., 1995), hundreds of different WD repeat proteins have been identified (Neer and Smith, 2000). These WD repeat proteins share a common beta-propeller structure composed of several four-stranded anti-parallel beta sheets/blades (Li and Roberts, 2001; Smith, 2008). Each WD repeat sequence encodes a structure of four beta anti-parallel strands but the repeat structure is not equivalent to a single blade. Rather, each repeat consists the first three strands of one blade and the last strand of the previous blade (Fig. 2.14e; Li and Roberts, 2001; Smith, 2008). In this way, the molecule can be more stabilized (Neer *et al*., 1994). These repeats also have a high percentage of hydrophobic residues to form the contact surface between the blades (Neer and Smith, 2000). The WD domain structure has two ends: the narrower end is often called the top, and wider end is called the bottom. In addition, there is a central tunnel that varies in shape and diameter based on the number of blades included (Neer and Smith, 2000).

The crystal structure of WDR5 is usually studied together with its associated complexes. Like other WD40 repeat proteins, WDR5 exhibits a typical propeller-like structure with seven blades and a narrow central tunnel (Han *et al*., 2006). Below I summarize important structures of WDR5 that interact with different proteins and complexes.


a
4
3
5
2
6
N
C
1
7
H3
N
C


b
MLL1
90°


c
RbBP5 peptide
4
5
3
6
2
7
1
H3
RbBP5


d
KANSL1
WDR5
Arg592
Val413
Leu411
408
KANSL2
416


e

Figure 2.14. Structure of WDR5. (a). Top and side views of crystal structure of WDR5 bound to N terminus of H3. H3 is bound to top face (the narrower end) of the β propeller of WDR5 (Taken from Couture *et al*., 2006 with permission). (b). Top and side views of crystal structure of WDR5 bound to MLL1. Similar to H3, MLL1 is also bound to top face of the β propeller of WDR5. (Taken from Song and Kingston, 2008 with permission) (c). Top and side views of crystal structure of WDR5 bound to RBBP5. RBBP5 is bound to the bottom face (the wider end) of the β propeller, (opposite site of H3 binding site). (d). Side view of crystal structure of WDR5 bound to KANSL1 and KANSL2. KANSL1 is bound to top face the β propeller of WDR5 and KANSL2 is bound to the bottom face (Taken from Dias *et al*., 2014 with permission). (e). WD repeat proteins share a common beta-propeller structure composed of several anti-parallel beta sheets/blades. The four-stranded anti-parallel beta sheets encoded by the same WD repeat sequence is marked in the same color. Each blade is composed of the first three parallel sheets of one WD repeat and the last parallel sheet of the previous WD repeat (Adapted from Valeyev *et al.,* 2008).

Recognition of Histone 3 - The top face (the narrower end) of the β propeller of WDR5 can recognize H3 protein, specifically the first three amino acids, Ala-Arg-Thr (ART, immediately preceding Lys4) of H3 (Fig. 2.14a, Han *et al*., 2006). There are several key amino acids in WDR5 involved in the binding to H3. The carboxylate side chain of Asp107 in WDR5 hydrogen bonds to the free amino group of Ala1 in H3. Ser91 and Cys261 in WDR5 also form a hydrogen bond to the backbone amide and carbonyl oxygen of Arg2 in H3 peptide. Tyr131 in WDR5 interacts with Ala1 in H3 via van der Waals contact (Han *et al*., 2006 and Couture *et al*., 2006). WDR5 have varied binding capacities to H3K4 with different degrees of methylation (Wysocka *et al*., 2005). Although the methyl groups of H3K4me2 are located on the surface of WDR5 with no obvious contact with WDR5, Han and colleagues believed that H3K4 dimethylation (H3K4me2) has the strongest binding affinity to WDR5 because the two methyl groups of Lys4 can form an extra pair of nonconventional hydrogen bonds with the carboxylate oxygen of Glu322 in WDR5. Compared with H3K4me2, H3K4 monomethylation (H3K4me1) has only one nonconventional hydrogen bond with Glu322, causing reduced interaction with

WDR5. The extra methyl group in H3K4me3 causes loss of the hydrogen bond, resulting in weaker affinity with WDR5 (Han *et al*., 2006 and Schuetz *et al*., 2006).

MLL1-WDR5-RBBP5 complex - As mentioned earlier, WDR5 together with RBBP5 and MLL can form an H3K4 methytransferase complex. MLL proteins can bind WDR5 through a Win motif. More detailed structure analysis indicated that Win motif of MLL1 binds WDR5 at the top face of the β propeller, which is the same binding pocket for H3 (Fig. 2.14b, Patel *et al*., 2008; Song and Kingston, 2008). There are several key amino acids in MLL1 Win motif that interact with WDR5. Gly3762 and Ser3763 of the MLL1 form multiple hydrogen bond and van der Waals interactions with Ala47, Ala65, Gly89, Ile90, and Asp107 in WDR5. The side chain of Arg3765 of MLL1 inserts into the central tunnel of WDR5 and form hydrogen bond and hydrophobic interactions with Ser91, Phe133, Ser175, Ser218, Cys261, Phe263, and Ile305 in WDR5. In addition, the Arg3765 is positioned between Phe133 and Phe263 of WDR5. These facts make Arg3765 of MLL1 the key amino acid for WDR5 binding. Ala47 and Tyr260 in WDR5 interacts with Ala3766 of MLL1 and Tyr-260 in WDR5 makes van der Waals contacts with Ala3766, Glu3767, and Val3768 in MLL1. His3769 of MLL1 interacts with WDR5 residues: Lys259, Phe149, Asp172, Pro173, and Tyr191 (Patel *et al*., 2008; Song and Kingston, 2008).

RBBP5 consists of an N-terminal β-propeller domain and a C-terminal hinge region. RBBP5 binds to WDR5 through its hinge region and more precisely, the segment located between residues 371 and 380 (Odho *et al*., 2010; Avdic *et al*, 2011). RBBP5 binds to a V-shaped cleft formed by the junction of blades 5 and 6 of WDR5 at its bottom face (the wider end) of the β propeller, which is on the opposite site of H3 and MLL1 binding sites (Fig. 2.14c). Similar to the binding of H3 and MLL1, WDR5 binds RBBP5 by maintaining several hydrogen

bonds and van der Waals interactions. The Gln289 side chain in WDR5 makes van der Waals contacts with Glu374 in RBBP5. WDR5 can form two hydrophobic pockets for RBBP5 binding. Pocket 1 is created by amino acid Leu249, Tyr228, Leu240 and Lys250 in WDR5 to make hydrophobic interactions with Val375 in RBBP5. And pocket 2 consists of Phe-266, Val-268, and Leu-288 in WDR5 to bind Val377 in RBBP5. Val 375 in RBBP5 also forms two hydrogen bonds with Gln289 side chain in WDR5. Asp376 and Ser379 in RBBP5 all form hydrogen bonds with Asn225 in WDR5 and WDR5′s Pro224 carbonyl group hydrogen bonds to Thr378 in RBBP5. There is also interaction between the RBBP5 Val380 and the WDR5 Arg181 (Odho *et al*., 2010; Avdic *et al*, 2011).

KANSL1/WDR5/KANSL2 complex - The nonspecific lethal (NSL) complex is a distinct multi-protein complex. Together with the KAT MOF (male absent on the first), they form an evolutionarily conserved KAT complex (Dias *et al*., 2014). WDR5, KANSL1 and KANSL2 are all subunits of NSL complex and both KANSL1 and KANSL2 can interact with WDR5 (Zhao *et al*., 2013; Dias *et al*., 2014). Interestingly, KANSL1 and KANSL2 binds to opposite sides of WDR5 β propeller like MLL1 and RBBP5 (Fig. 2.14d, Dias *et al*., 2014). KANSL1 interacts with WDR5 by a region centered on Arg592 in a similar manner to MLL proteins and H3. Arg592 in KANSL1 inserts into the central tunnel of the β propeller of WDR5 and form hydrogen bonds with Phe133, Phe263, Ser91, and Cys261 in WDR5. Thr587 and Val589 in KANSL1 form additional hydrogen bonds with Lys67, Ala65, and Gly89 in WDR5. Arg594 in KANSL1 makes a salt bridge with Asp107 in WDR5 and Val596 in KANSL1 makes a hydrophobic interaction with Tyr191, Pro216, and Leu234 in WDR5. Mutation on Arg592 in KANSL1 lead to unstable interaction with WDR5 and has resulted in protein degradation (Dias *et al*., 2014). KANSL2, like RBBP5, binds to the bottom face of the β propeller of WDR5.

Leu411, Val413, and Val414 of KANSL2 insert into a hydrophobic cleft between blades 5 and 6 of the WDR5 domain. The mutation of either Leu411 or Val413 in KANSL2 is sufficient to disrupt its interaction with WDR5 (Dias *et al*., 2014).

LncRNA binding interface on WDR5 - LncRNA HOTTIP and RBBP5 have shared binding cleft on WDR5 - the cleft between blades 5 and 6. Mutations of WDR5 on Tyr228, Leu240, Lys250, and Phe266 can significantly reduce its ability to bind HOTTIP. And among these mutations, F266A only affected its binding to HOTTIP without any defects in binding of MLL1 and RBBP5 (Yang *et al*., 2014).

2.3.2 Biological functions of WDR5

2.3.2.1 WDR5 is an important subunit in several protein complexes

WDR5 is required as a scaffold for multi-protein complex interactions. It plays a role in facilitating the assembly of many chromatin-modifying complexes. The function of WDR5 as a subunit of the Histone methyltransferase complex MLL/WDR5/RBBP5 was first discovered in 2005 (Wysocka *et al*., 2005) and is now becoming the most well studied function of WDR5. WDR5 is important for the recognition of H3 and for linking the methytransferase MLL to the rest of the WRAD complex. WDR5 is specifically and directly associated with methylated H3K4 and knockdown of WDR5 in human cells affects global H3K4 methylation levels (Wysocka *et al*., 2005). Together with the SET1-family complex, they achieve transcriptional activation via methylation of H3K4.

WDR5 is also a subunit of the histone acetyltransferase complex MOF/NSL. MOF is a member of the MYST family of histone acetyltransferases and is important in DNA damage repair and the maintenance of cell cycle and genomic stability (Zhao *et al*., 2013). MOF/NSL

complex can acetylate H4 on K16, K5, and K8 (Zhao *et al*., 2013). MOF directly interacts with KANSL1. WDR5 interacts with both KANSL1 and KANSL2 and brings the MOF/NSL complex together (Dias *et al*., 2014). In *Drosophila*, a point mutation of NSL1 that disrupts its interaction with WDS (WDR5 in *Drosophila*) is entirely lethal to females and partially lethal to males (Dias *et al*., 2014).

Since histone methylation on H3K4 and histone acetylation are all gene activation marks and both MLL/SET complex and MOF/NSL complex share a common WDR5 subunit, a question that needs addressing: is there any crosstalk between these two modifications? Indeed, there is a crosstalk between MOF-mediated H4K16 acetylation and MLL/SET-mediated H3K4 methylation. Mutation or knockdown of MOF led to a reduced global histone H3K4 di-methylation level in human cells while overexpression of MOF promoted H3K4 mono-, di- and tri-methylation. In contrast, the knockdown of RBBP5 only affected global H3K4 methylation level but had less effect on H4K16 acetylation. These findings indicate that MOF/NSL mediated H4 acetylation contributes to H3K4me2 by MLL/SET complexes (Zhao *et al*., 2013). More interestingly, WDR5 interacts with PKN1-mediated histone H3 threonine 11 phosphorylation (H3T11P), which in turn recruits MLL/SET complex leading to H3K4me3 occupancy (Kim *et al*., 2014). Since WDR5 is also a subunit of MOF/NSL complex, it is highly possible that H3T11P can also recruit MOF/NSL complex leading to H4 acetylation.

WDR5 is also reported to be a component of a CHD8 (chromodomain, helicase, DNA-binding) complex which is a group of ATP-dependent chromatin remodeling enzymes critical for the regulation of chromatin structure by modulating the contacts between histones and DNA (Thompson *et al*., 2008).

2.3.2.2 Role of WDR5 in regulation of *Hox* gene expression

As a component of multiple histone modification complexes, WDR5 plays an important role in gene regulation. *Hox* genes are well known targets of MLL/SET complex. Knockdown of WDR5 led to decreased H3K4me3 occupancy over the promoter regions of *HOXA9* and *HOXC8* and resulted in decreased expression of *HOXA9* and *HOXC8* (Wysocka *et al*., 2005). *Hoxa2* and *Hoxa10* are also targets of MLL/SET complex. The arginine methyltransferase PRMT6 (protein arginine methyltransferase 6) mediated H3R2 methylation can inhibit recruitment of WDR5 and H3K4 trimethylation and overexpression of PRMT6 resulted in decreased *Hoxa2* and *Hoxa10* expression (Hyllus *et al*., 2007). The lncRNA *HOTTIP* interacts with WDR5 and induces H3K4me3 on 5' HOXA genes. Knocking down of WDR5 greatly abolished transcription of *HOTTIP* and 5′ HOXA genes in primary foreskin fibroblast cells (Wang *et al*., 2011).

2.3.2.3 WDR5 is important in cell differentiation and development

WDR5 is also important for bone growth and skeletal development. WDR5 is expressed in osteoblasts, osteocytes as well as growth plate chondrocytes (Gori *et al*., 2001). The overexpression of *Wdr5* in a murine prechondroblastic cell line led to increased cAMP production and parathyroid hormone (PTH) binding as well as accelerated formation of mineralized nodules, which are all signs of increased osteoblastic differentiation (Gori *et al*., 2001). The persistent acceleration of osteoblast differentiation caused by overexpression of WDR5 was confirmed by postnatal analyses of transgenic mice and suggested that this effect is associated with the activation of the canonical Wnt pathway (Gori *et al*., 2006). *Wdr5* suppression resulted in significantly decreased alkaline phosphatase activity, Runx2 and osteocalcin expression, and absence of mineralized matrix formation, which are all indicators of

inhibited osteoblast differentiation (Zhu *et al*., 2008). WDR5 can also affect chondrocyte differentiation by modulating the expression of Twist-1 and fibroblast growth factor 18 (FGF18) (Gori *et al*., 2009). Down regulation of *Wdr5* in developing chick limb confirmed that *Wdr5* is required *in vivo* for endochondral bone development (Zhu *et al*., 2010). The down regulation of *Wdr5* resulted in impaired chondrocyte maturation, delayed endochondral bone development and shortened skeletal elements (Zhu *et al*., 2010). In ameloblastoma, abnormal bone resorption and bone formation are observed. Patients with ameloblastoma showed down regulated WNT-related bone-forming genes including WDR5 and Runx2, further supporting that WDR5 is important in bone formation (Sathi *et al*., 2012).

The transcription and regulatory networks in pluripotent ES cells are complicated and critical for the maintenance of ES cell differentiation and self-renewal and WDR5 is critical in maintaining epigenetic status in these networks. The expression of WDR5 has been found to be upregulated during the formation of induced pluripotent stem (iPS) cells and downregulated upon ES cell differentiation (Ang *et al*., 2011). Also higher levels of *Wdr5* expression and H3K4me3 occupancy were observed in ES cells compared to somatic cells and tissues, suggesting a role for *Wdr5* in ES and iPS cell maintenance (Ang *et al*., 2011). The knockdown of *Wdr5* induced changes in ES cell morphology and resulted in loss of self-renewal. Furthermore, the depletion of pluripotency transcription factor Oct4 led to down regulated *Wdr5* expression and decreased global H3K4me3 occupancy (Ang *et al*., 2011). In addition, the binding of lncRNA and WDR5 are important for ES cell pluripotency and differentiation. The binding of lncRNAs appears to be able to stabilize WDR5 protein and several ES cell lncRNAs that can bind WDR5 have been shown to be essential for maintaining ES cell fate (Yang *et al*., 2014). In mouse ES cells, expression of *LincRNA1230* can block the localization of Wdr5 at the promoters

of early neural genes and inhibit the cells from adopting a neural cell fate (Wang *et al*., 2016). As a component of KAT complex MOF/NSL, WDR5 is also important in facilitating the maintenance of H4K16ac in ES cells (Li *et al*., 2012; Mu *et al*., 2015).

2.3.2.4 Essential role of WDR5 in antiviral signaling pathway

Viral infection can cause the assembling of a virus-induced complex on the outer membrane of mitochondria and activate NF-ĸB and IRF3 to induce type I interferons (IFNs) (Wang *et al*., 2010). WDR5 is associated with virus-induced signaling adaptor (VISA), the key protein for assembling the virus-induced complex, and is translocated from nucleus to the mitochondria outer membrane. Knockdown of WDR5 disrupted the formation of VISA associated virus-induced complex and inhibited the trigger of NF-ĸB and IRF3 expression, suggesting that WDR5 plays an important role in cellular antiviral response (Wang *et al*., 2010).

2.4 Sumoylation and its role in nuclear protein translocation

Sumoylation is a posttranslational modification by an approximately 10 kDa Small Ubiquitin-like Modifier (SUMO) proteins. This family of small proteins can modify the function of their target protein by covalently attaching to and detaching from them. Three enzymes are involved in sumo conjugation process: an E1 activating enzyme, an E2 conjugating enzyme and an E3 sumo ligase. Sumoylation is involved in several important cellular processes, including cell cycle, subcellular transportation, regulation of transcription and protein stability (reviewed by Hay, 2005; Kumar and Zhang, 2015).

Among all the important functions of sumoylation, I have focused on its role in protein translocation. SUMOylation has emerged as a major mediator of nucleo-cytoplasmic

translocations of proteins. RanGAP1, the first SUMO substrate being identified, is a cytoplasmic protein (Matunis *et al*, 1996). The sumoylated RanGAP1 can activate the small GTPase protein Ran, a key player in the nuclear pore complex (NPC) that controls the nucleocytoplasmic trafficking of proteins (Mahajan *et al*., 1997). In HEK293 cells, five sumoylation sites have been identified on CDC73 protein yet only the sumoylation on K136 is located within the nuclear localization sequence and is responsible for the nuclear localization of CDC73 (Lamoliatte *et al*., 2014). The serine/threonine kinase GSK3β is a protein important in neuronal development. GSK3β can be sumoylated and is present in both cytoplasm and nucleus under normal conditions. A mutation of GSK3β that inhibit sumoylation resulted in exclusion of GSK3β in the nucleus (Eun Jeoung *et al*., 2008; Berndt *et al*., 2012). Three lysine residues of the IGF1R can be SUMO-1 sumoylated and the mutations of these SUMO-1 sites prohibited the translocation of IGF1R into the nucleus (Sehat *et al*., 2010). Sumoylation can also translocate proteins from nucleus to cytoplasm, examples include MEK1 and TEL. The cytoplasmic localization of both proteins are dependent on their sumoylation and disruption of sumoylation prevents the nuclear export of both proteins (Sobko *et al*., 2002; Wood *et al*., 2003). SUMO activation enzyme (SAE) can rapidly shuttle in and out of nucleus when it is not sumoylated at the C-terminus of SAE2 subunit. Sumoylation at the C-terminus of SAE2 subunit prevents the protein from being exported out from the nucleus (Truong *et al*., 2012).

CHAPTER 3

3. Goal of the project

3.1 Hypotheses to be tested：

Hoxa2 gene has been found to play a very important role in palate development in mice (Smith *et al*., 2009). DNA methylation is an epigenetic pathway that can regulate gene expression through the methylation of CpG islands on gene promoter. I first hypothesized that CpG islands exist on *Hoxa2* promoter and DNA methylation of the *Hoxa2* gene promoter will regulate *Hoxa2* expression in the developing mouse palate.

micorRNAs are small noncoding RNAs that can lead to gene silencing and have been found to regulate hox genes in embryo development. Yet only one miRNA has been reported to directly target *Hoxa2* gene. My second hypothesis is that specific miRNAs will directly bind *Hoxa2* 3' UTR and down regulate *Hoxa2* expression at both transcriptional and translational level.

Hotairm1 is a lncRNA identified in human that is transcribed between *Hoxa1* and *Hoxa2* and can activate the expression of *Hoxa1* and *Hoxa2*. The regulatory mechanisms of *Hotairm1* remain unstudied. My second hypothesis is that a noncoding transcript exists in the mouse genome between *Hoxa1* and *Hoxa2* genes and that can regulate their expression. This lncRNA will regulate *Hoxa1* and *Hoxa2* genes expression via histone modification. WDR5 is sumolylated and this modification is important in lncRNA induced histone methylation of *Hoxa1* and *Hoxa2* genes.

3.2 Objectives:

To test my first hypothesis, my objectives are to investigate the DNA methylation status of *Hoxa2* promoter in developing mouse palate and to determine whether *Hoxa2* expression will be regulate by DNA methylation in this process.

To test my second hypothesis, my objectives are to determine if specific miRNA binding sites are present on 3' UTR of *Hoxa2* mRNA and to determine if the miRNA's identified from above, are expressed in the developing mouse palate and whether this correlates with *Hoxa2* expression.

To test my third hypothesis, my objectives are to characterize a lncRNA transcribed between *Hoxa1* and *Hoxa2* and determine whether this lncRNA plays a role in regulating *Hoxa1* and *Hoxa2* expression. To identify the histone modification complex that interacts with the lncRNA. To investigate whether WDR5 is sumolylated and whether this modification plays a role in the function of WDR5.

CHAPTER 4

4. Materials and methods

4.1 Materials

4.1.1 Chemicals and reagents

Chemicals, reagents, kits, antibodies, miRNA mimic and siRNA, primers, vectors, enzymes and equipments used in the materials and methods section are listed below.

Table 4.1. List of chemicals, reagents and equipments.

Item and Catalog Number	Supplier
Chemical and Reagent	
1Kb Plus DNA Ladder (10787-018)	Invitrogen
Acrylamide:bis (1610125)	Bio-Red Laboratories, Ltd.
Agar (214010)	BD Biosciences
Agarose (BP1356-100)	Fisher Scientific, Co.
All trans retinoic acid (R2625)	Sigma-Aldrich, Co.
Ammonium persulfate (A9164)	Sigma-Aldrich, Co.
Ampicillin (A1066)	Sigma-Aldrich, Co.
β-mercaptoethanol (1610710)	Bio-Red Laboratories, Ltd.
Bovine serum albumin (BSA) (05471)	Sigma-Aldrich, Co.
Bovine calf serum (BCS) (SH30073.03)	HyClone Laboratories, Inc.
Coomassie blue (B0149)	Sigma-Aldrich, Co.
Denhardt's solution (D2532)	Sigma-Aldrich, Co.
DMSO (317275-500ML)	Millipore Canada, Ltd.
dNTPs (R0193)	Thermo Fisher Scientific, Co.
Dulbecco's Modified Eagle Medium (DMEM) (SH30022.01)	HyClone Laboratories, Inc.
EDTA (EDS-100G)	Sigma-Aldrich, Co.
EGTA (E3889)	Sigma-Aldrich, Co.
Ethanol (P016EAAN)	GreenField Specialty Alcohols, Inc.
Fetal bovine serum (FBS) (F1051)	Sigma-Aldrich, Co.
Fisher BioReagents EZ-Run Prestained *Rec* Protein Ladder (BP3603500)	Fisher Scientific, Co.
Glutathione Agarose (Pierce) (16100)	Thermo Fisher Scientific, Co.
Glycerol (BDH1172-4LG)	BDH Inc.

Glycine (BP381-1)	Fisher Scientific, Co.
HEPES (H3784)	Sigma-Aldrich, Co.
HiperFect transfection reagent (301705)	Qiagen Inc.
Kanamycin sulfate (60615)	Sigma-Aldrich, Co.
Lipofectamine 2000 (11668019)	Invitrogen
N-ethylmaleimide (NEM) (sc-202719)	Santa Cruz Biotechnology, Inc.
Methanol (A452-4)	Fisher Scientific, Co.
Formaldehyde (28908)	Sigma-Aldrich, Co.
Precision Plus Protein Kaleidoscope Protein Standards (1610375EDU)	Bio-Red Laboratories, Ltd.
Protease inhibitor cocktail (P8340)	Thermo Fisher Scientific, Co.
Protein A agarose (sc-2001)	Santa Cruz Biotechnology, Inc.
PVDF transfer membrane (88520)	Thermo Fisher Scientific, Co.
Streptavidin ultralink resin (Pierce) (20347)	Thermo Fisher Scientific, Co.
Sodium chloride (NaCl) (S271-3)	Fisher Scientific, Co.
Sodium dodecyl sulfate (SDS) (436143)	Sigma-Aldrich, Co.
Sodium hydrogencarbonate ($NaHCO_3$) (236527)	Sigma-Aldrich, Co.
Sodium pyruvate (SH30239.01)	HyClone Laboratories, Inc.
Spermidine (S2626)	Sigma-Aldrich, Co.
Spermine (S4264)	Sigma-Aldrich, Co.
Subcloning Efficiency™ DH5α™ Competent Cells (18265-017)	Invitrogen
Sucrose (S5-3)	Sigma-Aldrich, Co.
SUPERase In™ RNase Inhibitor (AM2694)	Ambion, Inc.
SYBR Green PCR Master Mix (4472908)	Applied Biosystems
Taqman universal master mix (4370048)	Applied Biosystems
TEMED (T7024)	Sigma-Aldrich, Co.
Tris (BP153-1)	Fisher Scientific, Co.
Triton X-100 (T9284-500ML)	Sigma-Aldrich, Co.
Tween 20 (BP337-500)	Thermo Fisher Scientific, Co.
Tryptone (BP1421-500)	Fisher Scientific, Co.
X-gal (CAS7240-90-6)	Fisher Scientific, Co.
Yeast extract (212750)	BD Biosciences
Commercially Available Kits	
Clarity Western ECL Blotting Substrate Kit (1705061)	Bio-Red Laboratories, Ltd.
EpiTect MSP kit (59305)	Qiagen, Inc.
EZ DNA methylation kit (D5001)	Zymo Research, Co.
GeneJET PCR Purification kit (K0701)	Thermo Fisher Scientific, Co.
GeneJET Plasmid Miniprep Kit (K0503)	Thermo Fisher Scientific, Co.
IsHyb In Situ Hybridization Kit (K2191050)	BioChain Institute, lnc.
LucPair miR Duo-Luciferase Assay Kit (LPFR-M010)	GeneCopoeia

miScript II RT Kit (218161)	Qiagen, Inc.
miScript SYBR Green PCR Kit (218073)	Qiagen, Inc.
PureLink Genomic DNA Mini Kit (K1820-01)	Invitrogen
QIAquick Gel Extraction kit (28704)	Qiagen, Inc.
QuantiTect Reverse Transcription Kit (205311)	Qiagen, Inc.
QuikChange Lightening Site-Directed Mutagenesis Kit (210518)	Agilent Technologies
RNeasy Protect Mini Kit (74124)	Qiagen, Inc.
Antibodies	
Anti-Digoxigenin-fluorescein, Fab fragments (11207741910)	Roche Diagnostics
GAPDH antibody (G-9) (sc-365062)	Santa Cruz Biotechnology, Inc.
Goat Anti-Mouse IgG (H+L) HRP Conjugate (1706516)	Bio-Red Laboratories, Ltd.
Goat Anti-Rabbit IgG (H+L) HRP Conjugate (1706515)	Bio-Red Laboratories, Ltd.
Histone H3 antibody (ab1791)	Abcam, Inc.
Histone H3 (tri methyl K4) antibody (ab8580)	Abcam, Inc.
Histone H3 (tri methyl K27) antibody (ab6002)	Abcam, Inc.
Normal mouse IgG (sc-2025)	Santa Cruz Biotechnology, Inc.
Normal mouse IgM (sc-3881)	Santa Cruz Biotechnology, Inc.
Normal rabbit IgG (sc-2027)	Santa Cruz Biotechnology, Inc.
SUMO1 21C7 antibody	Developmental Studies Hybridoma Bank
WDR5 antibody (A-6) (sc-393116)	Santa Cruz Biotechnology, Inc.
WDR5 antibody (ab22512)	Abcam, Inc.
Commercially Available miRNA mimic and siRNA	
AllStars Negative Control siRNA (SI03650318)	Qiagen, Inc.
Syn-mmu-miR-376c-3p miScript miRNA Mimic (MSY0003183)	Qiagen, Inc.
Syn-mmu-miR-669b-5p miScript miRNA Mimic (MSY0003476)	Qiagen, Inc.
Commercially Available Primers	
FAM-labeled Hoxa2 Taqman primer	Applied Biosystems
Hs_RNU6-2_11 miScript Primer Assay (MS00033740)	Qiagen, Inc.
Mm_miR-19a_1 miScript Primer Assay (MS00001302)	Qiagen, Inc.
Mm_miR-298_1 miScript Primer Assay (MS00002016)	Qiagen, Inc.
Mm_miR-376c_1 miScript Primer Assay	Qiagen, Inc.

(MS00002268)	
Mm_miR-431_1 miScript Primer Assay (MS00002373)	Qiagen, Inc.
Mm_miR-669b_1 miScript Primer Assay (MS00002716)	Qiagen, Inc.
Mm_miR-878-3p_1 miScript Primer Assay (MS00012789)	Qiagen, Inc.
Prime Time qPCR Primers Hoxa1 (Mm.PT.51.10819147.g)	IDT Integrated DNA Technologies, lnc.
Prime Time qPCR Primers Hoxa2 (Mm.PT.51.7231226)	IDT Integrated DNA Technologies, lnc.
VIC-labled β-actin Taqman primer	Applied Biosystems
Vectors	
pEZX-MT01 (CmiT000001-MT01)	GeneCopoeia
pEZX-MT01-Hoxa2 (MmiT029383-MT01)	GeneCopoeia
pGEM-T Easy Vector System I (PR-A1360)	Promega, Co.
pGEX-6p-1 (28-9546-48)	GE Healthcare
Enzymes	
BamH I (FastDigest) (FD0054)	Thermo Fisher Scientific, Co.
EconoTaq DNA polymerase (30031-2)	Lucigen Co.
EcoR I (FastDigest) (FD0274)	Thermo Fisher Scientific, Co.
Proteinase K (P2308)	Sigma-Aldrich, Co.
Equipment	
AccuSpin Micro 17	Thermo Fisher Scientific, Co.
Branson digital sonifier 250	Branson Ultrasonics
CKX41 Microscope	Olympus Canada, INC.
Isotemp Digital Block Dry Heater	Thermo Fisher Scientific, Co.
Milli-Q Advantage A10 Water Purification System	Millipore Canada, Ltd.
Monobloc Weighing Technology AB54-S	METTLER TOLEDO
Moxi Z Mini Automated Cell Counter	ORFLO Technologies
MyCycler Thermal Cycler	Bio-Red Laboratories, Ltd.
NanoVue Plus Spectrophotometer	GE Healthcare
pH 700 Benchtop Meter	OAKTON Instruments
PowerPac Basic Power Supply	Bio-Red Laboratories, Ltd.
Sorvall Legend X1 Centrifuge	Thermo Fisher Scientific, Co.
StepOne Real-Time PCR System	Applied Biosystems Canada
Syngene G:Box	Syngene
UV Transilluminator TM-10E	UVP
Vortex-Genie 2	VWR

Table 4.2. Names and addresses of distributors

Company	Location
Abcam, Inc.	Toronto, ON, Canada
Agilent Technologies	Santa Clara, CA, USA
Ambion, Inc.	Foster City, CA, USA
Applied Biosystems Canada	Streetsville, ON, Canada
Branson Ultrasonics	Danbury, CT, USA
BD Biosciences	Mississauga, ON, Canada
BDH Inc.	Toronto, ON, Canada
BioChain Institute, lnc.	Newark. CA, USA
Bio-Red Laboratories, Ltd.	Mississauga, ON, Canada
Developmental Studies Hybridoma Bank	Iowa City, IA, USA
Fisher Scientific, Co.	Ottawa, ON, Canada
GE Healthcare	Mississauga, ON, Canada
GeneCopoeia	Rockville, MD, USA
GreenField Specialty Alcohols, Inc.	Brampton ON, Canada
HyClone Laboratories, Inc.	Logan, UT, USA
IDT Integrated DNA Technologies, lnc.	Coralville, IA, USA
Invitrogen Canada, Inc.	Burlington, ON, Canada
Lucigen, Co.	Middleton, WI, USA
METTLER TOLEDO	Mississauga, ON, Canada
Millipore Canada, Ltd.	Etobicoke, ON, Canada
Santa Cruz Biotechnology, Inc.	Dallas, TX, USA
Sigma-Aldrich, Co.	Oakville, ON, Canada
Thermo Fisher Scientific, Co.	Waltham, MA, USA
OAKTON Instruments	Vernon Hills, IL, USA
Olympus Canada, INC.	Richmond Hill, ON, Canada
ORFLO Technologies	Ketchum, ID, USA
Promega, Co.	Madison, WI, USA
Qiagen, Inc.	Mississauga, ON, Canada
Roche Diagnostics	Laval, QC, Canada
Syngene	Cambridge, UK
UPV	Upland, CA, USA
VWR	Mississauga, ON, Canada
Zymo Research, Co.	Irvine, CA, USA

4.1.2 Cell lines

In this study, NIH 3T3 (ATCC CRL-1658) and EG7 (ATCC CRL-2113) cell lines were used. NIH 3T3 is a mouse embryonic fibroblast cell line that exhibits expression of both *Hoxa1*

and *Hoxa2*. In addition, *Hoxa2* promoter was demonstrated to be unmethylated in this cell line (X. Wang, PhD Thesis, 2013). EG7 cell line is derived from the murine T-cell lymphoma EL-4 transfected with cDNA for ovalbumin (OVA). Previous studies from our laboratory has shown *Hoxa2* promoter to be highly methylated with no *Hoxa2* expression in the EG7 cell line (X. Wang, PhD Thesis, 2013).

4.1.3 Animals

Female timed-pregnant CD-1 mice were obtained from Animal Resources Centre, University of Saskatchewan. All procedures were approved by the University Committee on Animal Care and Supply at the University of Saskatchewan.

4.2 Methods

4.2.1. Cell culture conditions

The mouse embryonic fibroblast NIH 3T3 cell line was cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% bovine calf serum (BCS) and 1% antibiotic/antimycotic (Ab/Am). The mouse lymphocyte EG7 cell line was cultured in DMEM supplemented with 10% FBS, 110 mg/L sodium pyruvate and 1% Ab/Am. Both cell cultures were incubated with 5% CO_2 and 100% relative humidity at 37° C.

4.2.2. DNA methylation analysis

4.2.2.1 Sodium bisulfite modification

Genomic DNA samples were extracted from NIH 3T3 cell line, EG7 cell line and E12 to E15 mouse palate tissue using PureLink® Genomic DNA Mini Kit (Invitrogen). Genomic DNA

was modified with bisulfite reagent for DNA methylation analysis using the EZ DNA methylation kit (Zymo Research, USA). According to the manufacturer's instructions, 2 μg of genomic DNA was modified per sample. NanoVue UV/visible Spectrophotometer (GE Healthcare) was used to determine the concentration and purity of bisulfite treated DNA samples. Bisulfite treated DNA samples were stored at -20°C.

4.2.2.2 Methylation specific PCR

After sodium bisulfite modification of genomic DNA, DNA methylation within promoter associated CpG islands was determined with methylation specific PCR (MSP) primers listed in Table 4.3. The principle used to design MSP primers are shown in Figure 4.1. Bisulfite treatment of DNA can convert unmethylated cytosine (C) into uracil (U) while the methylated C will remain as C. In this way the methylation status of a CpG island can be identified by its sequence after bisulfite conversion. MSP primers were designed to include at least one CpG site in the primer sequence and based on the sequence differences between methylated and unmethylated CpG island following bisulfite treatment. Two sets of primers were designed for each CpG island, the methylated primer set and the unmethylated primer set. In total, six pairs of MSP primers were designed to detect methylation status of the three CpG rich regions of the *Hoxa2* promoter (Table 4.3).

For the PCR amplification (a total of 50μl), 200 ng of bisulfite-modified DNA, 2 × EpiTect MSP master mix (Qiagen), 200 μM of each dNTP and 500 nM of each forward and reverse primer was used. The PCR cycle for region 1 and 3 (MPS-CpG1 and MPS-CpG3) started with a Taq activation step at 95°C for 10 min, followed by 35 cycles of 94°C for 15sec, 53°C for 30sec, 72°C for 30sec and a final extension step at 72°C for 10 min. The PCR cycle for region 2

(MPS-CpG2) had 1 cycle of 95°C for 10 min, followed by 35 cycles of 94°C for 15sec, 55°C for 30sec, 72°C for 30sec and a final extension step at 72°C for 10 min.

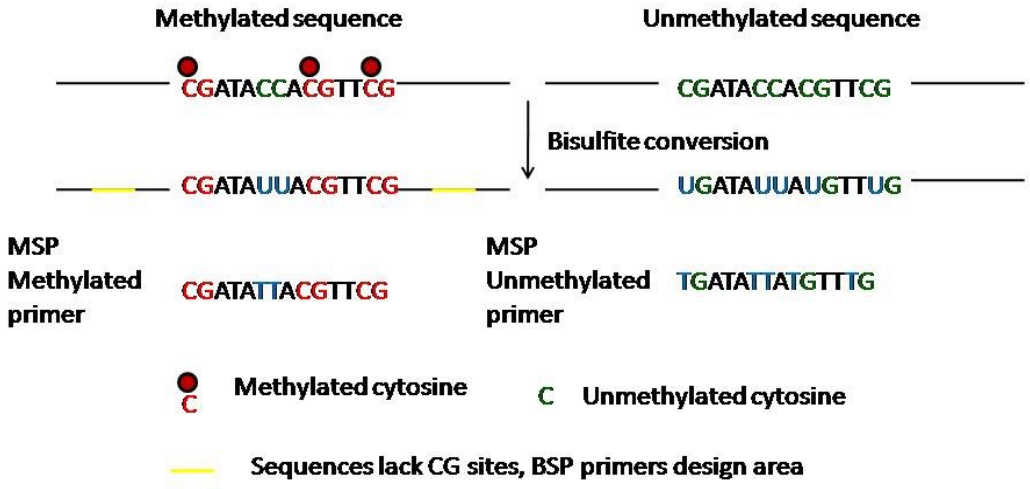


Figure 4.1. Bisulfite determination of DNA methylation. Under the treatment of bisulfite conversion, the methylated cytosine (C) in the genomic DNA will remain as C, while unmethylated C will be convered into uracil (U). Based on this, methylation specific primers (MSP) can be designed for methylated and unmethylated sequences. During PCR, U will be replaced with thymine (T). BSP primers are designed in the sequences that lack CG sites.

4.2.2.3 Bisulfite specific PCR and sequencing

After sodium bisulfite modification of genomic DNA, bisulfite specific PCR (BSP) amplification was carried out. BSP primers were designed in the sequences that lack CG sites so the CpG island, whether methylated or not, can be amplified by the same BPS primer and the sequence difference can be tested by DNA sequencing (Fig. 4.1). For the PCR amplification (a

total of 50µl), 200 ng of bisulfite-modified DNA, 10×EconoTaq buffer, 2.5 units of EconoTaq DNA polymerase (Invitrogene), 200 µM of each dNTP and 500 nM of each forward and reverse primer (Table 4.3) were used. The PCR cycle started with one step at 95°C for 3min, followed by 35 cycles of 95°C for 30sec, 55°C for 30sec, 72°C for 40sec and a final extension step at 72°C for 10 min. DNA samples were run on a 1% agarose gel and target bands were excised from the gel and purified using QIAquick Gel Extraction kit (Qiagen). The purified DNA samples were then cloned into pGEM® T-easy vector (Fig.4.2, Promega). The ligation system (10 µl) was set up with 5µl 2x Rapid Ligation Buffer, 1 µl pGEM® T-easy vector, 1 µl T4 DNA ligase and 3 µl gel purified PCR product (pGEM®-T Easy Vector System I, Promega). Ligation was carried at 10°C over night, followed by 5 min incubation at 70°C to inactivate T4 DNA ligase. Ligation product was then transformed into DH5α™ Competent Cells (Subcloning Efficiency™, Invitrogen). For each transformation, 5 µl ligation product was incubated with 50 µl DH5α™ competent cells for 40 min on ice and heat shocked at 42°C for 90 s. Cells were then incubated on ice for 2 min and recovered in 200 µl LB medium (10g Bacto-Tryptone, 5g Bacto-yeast extract and 10g NaCl in 1L medium, PH 7.4) for 30 min at 37°C. The cells were then spread in LB-ampicillin agar plates (10 ml LB per plate with 1% ampicillin and 1.5% agar) with 20mM IPTG and 80µg/ml X-gal for blue-white screening. The plates were incubated at 37°C overnight and each single white clone was picked and cultured in 4ml LB with 1% ampicillin overnight at 37°C. Cells were then collected and plasmids were purified with GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) as per manufacture's protocol. The plasmids containing BSP amplification fragments were sequence analyzed with an ABI PRISM^{TM} system using M13 forward primer at the National Research Council, Plant Biotechnology Institute, Saskatoon, SK.

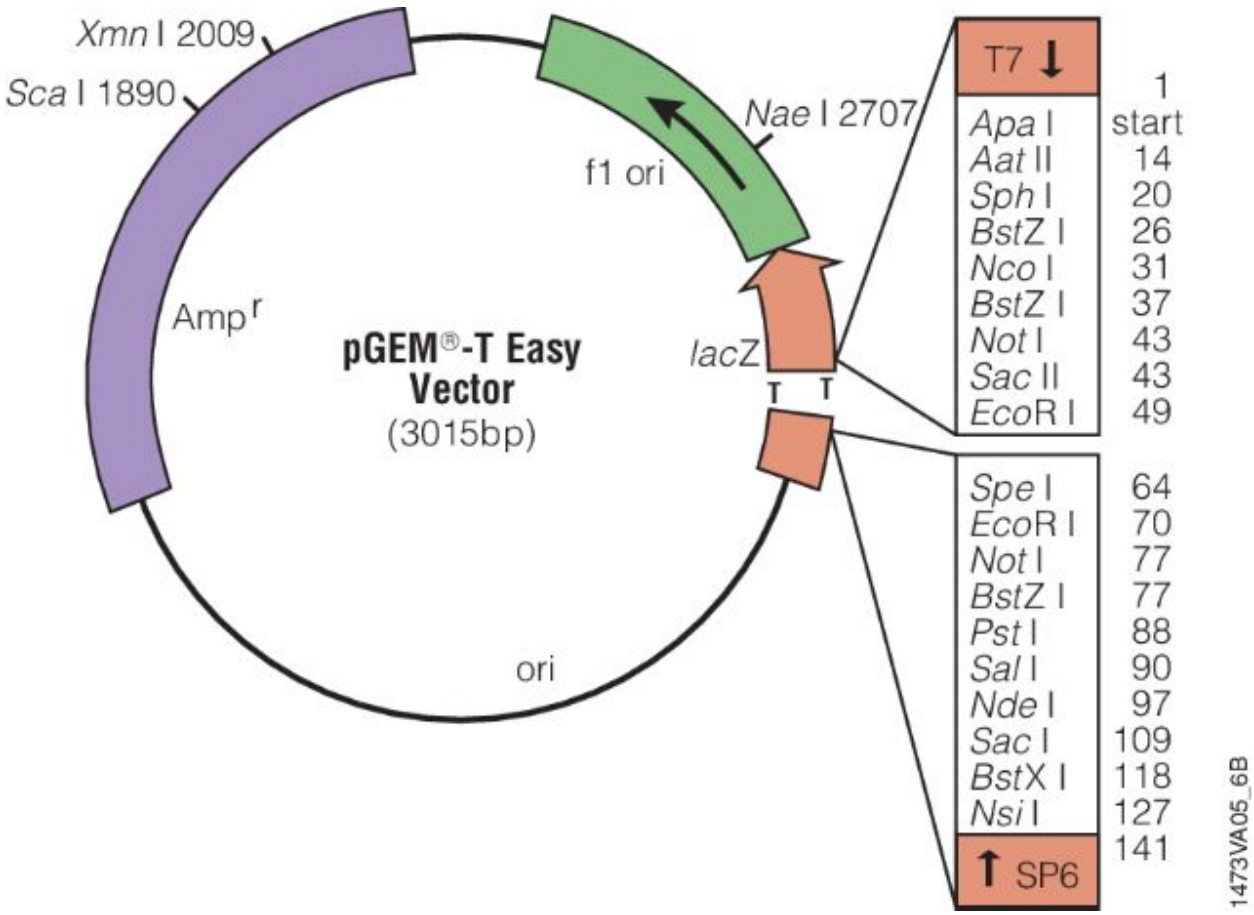


Figure 4.2. pGEM® T-easy vector map (Promega). pGEM®-T Easy Vector is a linearized vector with a single 3´-terminal thymidine (T) at both ends. PCR products generated by certain polymerases are adenine (A) tailed and can be ligated to pGEM®-T Easy Vector based on A-T pairing. pGEM®-T Easy Vector has the α-peptide coding region of the enzyme β-galactosidase (*lac*Z) which allows blue/white screening on indicator plates. The pGEM®-T Easy Vector also provides three single-enzyme digestions (*Eco*RI, *Bst*ZI and *Not*I) for release of the insert. Figure taken from Promega http://embed.widencdn.net/img/promega/ws62vit9cf/640px/1473VA05_6B.jpeg?u=7fvzhm.

4.2.3. Prediction and analysis of miRNA binding sites on mouse *Hoxa2* 3' UTR

The website http://www.microrna.org was used to predict miRNAs that may bind to mouse 3' UTR of *Hoxa2* gene, and comparisons made with the binding site present in mouse 3' UTR *Hoxa2* gene with other species including human, rat, chimpanzee and dog using Genbank (NCBI).

4.2.4. RNA Isolation and Reverse Transcription (RT)

For the mRNA study, total RNA was isolated from NIH 3T3 cells, mouse embryo palate shelves (E12 to E15) and mouse tissue samples (head, forelimb, hindlimb and tail) from E13 embryos using RNeasy Protect Mini Kit (Qiagen). Each sample was reverse transcribed to cDNA using QuantiTect® Reverse Transcription Kit (Qiagen). Genomic DNA (gDNA) elimination reaction was set up to a total volume of 14 µl with 2 µl 7x gDNA Wipeout Buffer, 500ng template RNA and incubated for 2 min at 42°C and placed on ice immediately. Reverse-transcription master mix was prepared with 1 µl Quantiscript Reverse Transcriptase, 4 µl 5x Quantiscript RT Buffer and 1 µl RT Primer Mix and added to gDNA eliminated RNA template from the previous step. The reverse transcription was then carried out by incubating the system at 42°C for 1 h and 3 min at 95°C to inactivate Quantiscript Reverse Transcriptase. cDNA samples were then stored at -20°C for further experiments. For miRNA study, total RNA including small RNA was isolated from NIH 3T3 cell culture and mouse embryo palate shelves (E12 to E15) using miRNeasy Mini Kit (Qiagen) and was reverse transcribed into cDNAs using miScript II RT Kit (Qiagen). A total volume of 20 µl reverse-transcription reaction containing 4 µl 5x miScript HiSpec Buffer, 2 µl 10x miScript Nucleics Mix, 2 µl miScript Reverse Transcriptase Mix and 500 ng template RNA was incubated for 60 min at 37°C followed by 5 min incubation at 95°C to inactivate miScript Reverse Transcriptase Mix. cDNA samples were stored at - 20°C for further experiments.

4.2.5. Quantitative real-time PCR (qRT-PCR)

In miRNA study, *Hoxa2* gene expression was detected using FAM (fluorescein amidite)-labeled Taqman gene expression primers (Applied Biosystems). VIC-labeled β-actin Taqman

primers were used as endogenous control. PCR reaction started with a 2 min initiation at 50°C and 10 min denaturation at 95°C, followed by 40 cycles with each cycle consisting of 15 s at 95°C and 70 s at 60°C (Smith *et al*., 2009). Mature miRNAs were detected with individual miRNA SYBR® green primers (Qiagen). Small nuclear RNA (snRNA) RNU6B was used as control gene of miRNA expression and was detected with RNU6B SYBR® green primers (Qiagen). Amplification system of 25µl was prepared with miScript SYBR® Green PCR Kit (Qiagen) with 25ng cDNA input and the reaction cycles were as follows: 15 min initial activation at 95°C, followed by 40 cycles of 15s at 94°C, 30s at 55°C and 30s at 70°C in 7300 ABI detection system. All reactions were run in replicates of 2, n=4.

For the lncRNA study, *Hoxa1, Hoxa2* and *mHotairm1* expressions were detected using primers as listed in Table 4.1 and Table 4.3. β-actin was detected using primers listed in Table 4.3 as control house keeping gene. SYBR® Green PCR Master Mix (Applied Biosystems) was used and the reaction cycles were as follows: 10 min initiation at 95°C, followed by 40 cycles of 15s at 95°C and 60s at 60°C in 7300 ABI detection system. All reactions were run in replicates of 2, n=3.

Relative quantity (RQ) value was determined by the $2^{(-\Delta\Delta Ct)}$ method described in the (Smith *et al*., 2009) and from Applied Biosystems User Bulletin (ABI PRISM 7700 Sequence Detection System, Applied Biosystems, 2001).

4.2.6. Western blot analysis

Protein samples were collected from NIH 3T3 cells and E13 *Hoxa2*$^{-/-}$ mice embryos with radioimmunoprecipitation assay buffer (RIPA buffer). Samples were then quantified using a Bio-Rad DC Protein Assay kit and denatured for 10 min with 2X loading buffer. Each sample was

electrophoresed in a 10% polyacrylamide-SDS gel and transferred to a PVDF transfer membrane (Thermo Fisher Scientific) using Xcell II Blot Module (Invitrogen) for wet transfer in 1X transfer buffer containing 12 mM Tris, 96 mM glycine and 20% methanol, for 1.5 h at 20V/150 mAmp.

In miRNA study, the blotting and detection of Hoxa2 protein was performed as described in Smith *et al*. (2009). Briefly, PVDF membranes were blocked in 3% skim milk (in PBS) at 4 °C for 1 h and then incubated with the rabbit anti-Hoxa2 antiserum (Hao *et al*., 1999) at a dilution of 1:1000 in 3% skim milk overnight at 4 °C. Membrane were then washed 3 x 10 min with 0.1% Tween-20 in PBS (PBST) and incubated for 1 h at room temperature with horse radish peroxidase (HRP) conjugated secondary goat anti-rabbit IgG (1:3000 in 5% skim milk, Bio-Rad). Visualization of proteins was performed by a Clarity™ Western ECL Blotting Substrate Kit (Bio-Rad) as per manufacture's instructions, followed by exposure of the membranes using Syngene imaging system (Syngene, Cambridge, UK). The same membranes were probed with GAPDH antibody as an internal control (sc-365062, 1:3000, Santa Cruz®) for 1h at room temperature and secondary anti-mouse IgG antibody (1:3000, Bio-Rad®) for 1h at room temperature.

In lncRNA study, to detect WDR5 protein, transfer membranes were blocked in 3% bovine serum albumin (BSA) for 1 h (in PBST) followed by overnight incubation with anti-WDR5 antibody (ab22512, 1:1000, Abcam®) at 4°C, and subsequent incubation with secondary anti-rabbit IgG antibody (1:3000, in 5% skimmed milk in PBST, Bio-Rad®) for 1 h at room temperature. Sumoylated protein was detected by overnight incubation with SUMO1 21C7 antibody (1:500, in 3% BSA, Developmental Studies Hybridoma Bank) at 4°C and secondary anti-mouse IgG antibody (1:3000, in 5% skimmed milk, Bio-Rad®) for 1h at room temperature.

4.2.7. MiR-669b and miR-376c over expression

NIH 3T3 cells (6 $\times 10^4$/well) were seeded in a 24-well plate in 0.5 ml culture medium. HiperFect Transfection Reagent, (6 µl, Qiagen®) and 300 ng of either miRNA-669b mimic, miR-376c mimic (Qiagen®) or control siRNA (Qiagen®) were applied to each well based on HiperFect transfection protocol (Qiagen®). Cells were then incubated under normal growth conditions and RNA samples were isolated after 24 h and 48 h incubation. Protein samples were collected after 24 h incubation.

4.2.8. Hoxa2-3'UTR mutation vector construction

The vector pEZX-MT01-*Hoxa2* (GeneCopoeia^{TM}) contained full length mouse *Hoxa2* 3'UTR. QuikChange Lightening Site-Directed Mutagenesis Kit (Agilent Technologies) was used to carryout site directed mutations on *Hoxa2* 3'UTR. pEZX-MT01-*Hoxa2mut1* was generated using forward primer 669b1F and reverse primer 669b1R (Table 4.3) with an AA-CG mutation (Fig. 5.5A). pEZX-MT01-*Hoxa2mut2* was generated using forward primer 669b2F and reverse primer 669b2R (Table 4.3) with an AA-CG mutation (Fig.5.5A). Mutation reaction was set up (50 µl) with 5 µl of 10x reaction buffer, 100ng pEZX-MT01-*Hoxa2* vector template, 125 ng of forward and reverse primers, 1 µl of dNTP mix, 1.5 µl of QuikSolution reagent and 1 µl of QuikChange Lightning Enzyme. The mutation PCR started with 2 min incubation at 95°C, followed by 18 cycles of 20s at 95°C, 10s at 60°C and 3 min at 68°C, and a 5 min extension at 68°C. The PCR products were then digested with 2 µl of *Dpn* I restriction enzyme at 37°C for 5 min to digest the parental vector DNA. Mutated vectors were then transformed into XL10-Gold ultracompetent cells as described in manufacturer's protocol (QuikChange Lightening Site-Directed Mutagenesis Kit, Agilent Technologies). Each single clone was picked and incubated in

4 ml LB medium with 1% Kanamycin over night at 37°C. Plasmid was purified from bacterial culture using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) as per manufacture's protocol and target mutation was confirmed via sequencing with an ABI PRISM^{TM} system at the National Research Council, Plant Biotechnology Institute, Saskatoon, SK. A double mutation vector pEZX-MT01-*Hoxa2mut1+2* containing both mutation1 and mutation2 was also generated.

4.2.9. Luciferase assay

The vector pEZX-MT01 (Genecopoeia) encoding *firefly* luciferase (hLuc) and *renilla* luciferase (hRLuc) was used as negative control in luciferase assay (Fig. 4.3A). NIH 3T3 cells (6 $\times 10^4$) were seeded in each well of a 24-well plate in 0.5 ml culture medium and cultured for 24 h. After 24 h incubation, either miR-669b mimic (Qiagen), miR-376c mimic (Qiagen) or control siRNA (Qiagen) was co-transfected with one of the following luciferase vectors: pEZX-MT01-*Hoxa2*, pEZX-MT01-*Hoxa2mut1*, pEZX-MT01-*Hoxa2mut2*, pEZX-MT01-*Hoxa2mut1+2* or blank pEZX-MT01 vector using lipofectamine 2000 (Invitrogen). For each transfection, 3 µl lipofectamine 2000 (Invitrogen) was incubated together with 300 ng miRNA mimic/siRNA control and 1 µg luciferase vector in 100 µl DMEM for 20 min and then added into cell cultures. The vector pEZX-MT01 (Genecopoeia) encoding hLuc and hRLuc was used as negative control in luciferase assay. After 24h post-transfection, cells were washed with PBS and lysed using LucPair miR Duo-Luciferase Assay Kit (GeneCopoeia). *Firefly* and *Renilla* luciferase activities were determined for each transfection using a luminometer. The *firefly* luciferase signals were normalized to the *Renilla* luciferase signal.

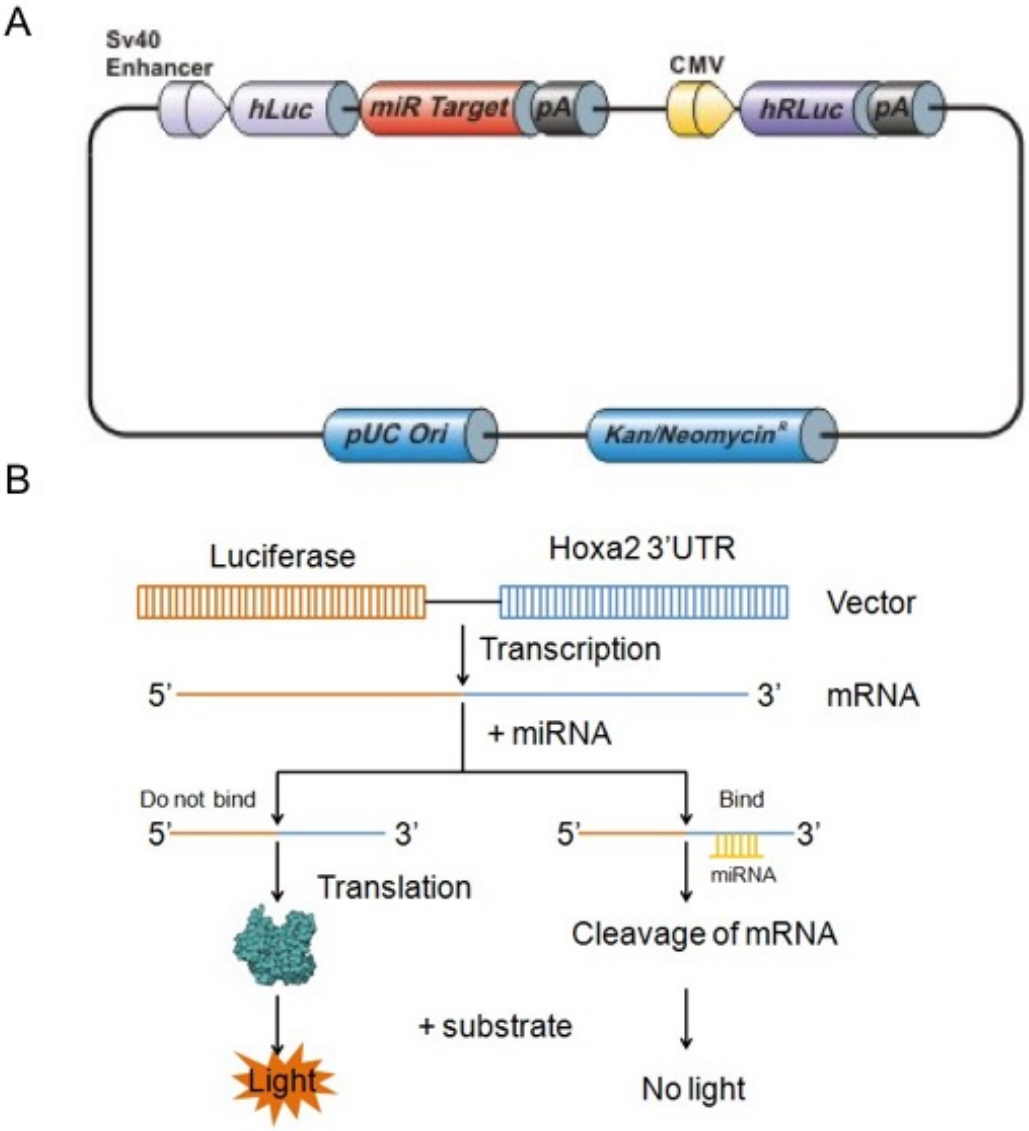


Figure 4.3. Mechanism of Duo-Luciferase assay. (A) pEZX-MT01 vector map. *Hoxa2* 3'UTR is inserted as miR target. pEZX-MT01 vector encodes *firefly* luciferase (hLuc) and *renilla* luciferase (hRLuc). hLuc mRNA is merged with *Hoxa2* 3'UTR and is used as an indicator of miRNA effects. hRluc is expressed independently and is used as an internal control. Vector map taken from Genecopoeia (B) Mechanism of luciferase assay. *Hoxa2* 3'UTR is attached to *firefly* luciferase mRNA. If *Hoxa2* 3'UTR is targeted by miRNAs, the entire mRNA may be cleaved and reduced luminescence will be detected.

4.2.10. Mouse *Hotairm1 (mHotairm1)* sequence analysis

Three pairs of overlapping primers (Mush 147F, Mush 147R, Mush 138F, Mush 138R, Mush 245F, Mush 245R) were designed (Fig. 4.4) and PCR were carried out using cDNA samples from NIH 3T3 cells. For the PCR amplification (a total of 50µl), 50 ng of cDNA, 10×EconoTaq buffer, 2.5 units of EconoTaq DNA polymerase (Invitrogen), 50 µM of each dNTP and 500 nM of each forward and reverse primer (Table 3.1) were used. The PCR cycle started with one cycle at 95°C for 3 min, followed by 30 cycles of 95°C for 30 sec, 52°C for 30 sec, 72°C for 40 sec and a final extension step at 72°C for 10 min. The PCR fragments were run on a 1.2% agarose gel and purified from the gel using QIAquick Gel Extraction kit (Qiagen). The DNA samples were then cloned into pGEM T-easy vector as described in 4.2.2.3 for subsequent sequence analysis with an ABI PRISM^{TM} system at the National Research Council, Plant Biotechnology Institute, Saskatoon, SK.

4.2.11. *MHotairm1* silencing in NIH 3T3 cells

NIH 3T3 cells (6×10^4) were seeded in each well of a 24-well plate in 0.5 ml culture medium. siRNA (150 ng of *hotairm1* siRNA or control siRNA) and 6 µl HiperFect Transfection Reagent (Qiagen) were incubated in 100 µl DMEM medium for 20 min and then added to each well. Cells were then incubated under normal growth conditions and RNA samples were isolated after 72 h incubation. SYBR green real-time PCR was carried out with *mHotairm1* primers mush 137F and mush138R (Table 4.3). *Hoxa1* and *Hoxa2* expression were detected using *Hoxa1* and *Hoxa2* SYBR primers (Table 4.1, Integrated DNA Technologies).

4.2.12. Effect of all-*trans* Retinoic acid (ATRA) on *mHotairm1* expression

NIH 3T3 cells were seeded in 100 mm cell culture dish and grown to 70% confluency. NIH 3T3 cells were treated with ATRA (10^{-6}M) for 24 h. RNA was isolated, reverse transcribed to cDNA, followed by real-time PCR to detect *mHotairm1*, *Hoxa1* and *Hoxa2* expression as indicated in 4.2.4 and 4.2.5. Protein samples were collected and WDR5 protein was detected using western blot analysis as indicated in 4.2.6.

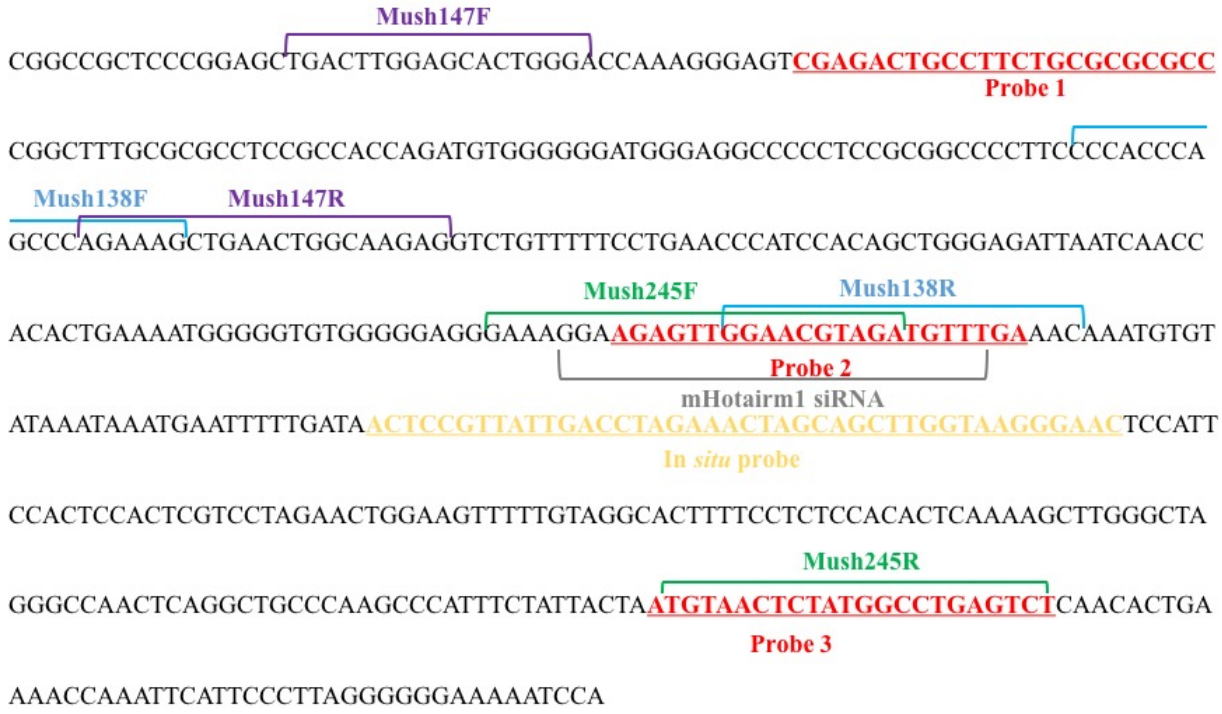


Figure 4.4. Primer, siRNA and probe design for *mHotairm1*. Three pairs of primers were designed to cover the predicted mouse *hotairm1* sequence. Violet: primer Mush 147F/R; Blue: primer Mush 138F/R; Green: primer Mush 245F/R. Primer names were marked above sequence. Three probes (Probe 1, 2 and 3, underlined sequences in red) were designed against *mHotairm1* for CHART experiment and Probe C is complimentary to Probe 2. CHART Probe names were marked below sequence. mHotairm1 siRNA is marked in grey and *mHotairm1* in *situ* probe is marked in yellow.

4.2.13. *In situ* hybridization histochemistry

Mouse head samples were dissected and immediately frozen in isopentane on dry ice. *In situ* hybridization histochemistry was carried out using IsHyb In Situ Hybridization (ISH) Kit (BioChain) as per manufacture's protocol. Frozen palatal sections (20 μm) were collected on slides and fixed immediately in 4% paraformaldehyde in PBS for 10 min. The samples were then rinsed with RNase free PBS and treated with proteinase K digestion (10 μg/ml in PBS) for 5 sec. When the slides were dry, pre-hybridization solution was added and slides were incubated for 3h at 50°C. Slides were then washed with 2x Saline Sodium Citrate (SSC) for 10 min at 45°C, 1.5x SSC for 10 min at 45°C and 0.2x SSC twice for 20 min at 37°C. Slides were blocked with 1x blocking solution for 1 h at room temperature. Labeled 100 ng of digoxigenin (DIG) tagged *mhotairm1* probe (Fig. 4.4, Table 4.3, IDT) or random control probe was added to the slides, respectively and incubated over night at 37°C. Slides were then rinsed with SSC buffer and PBS and blocked for 30 min at room temperature with 0.1% Triton 100, 1% sheep serum in PBS. The anti-DIG-fluorescein, Fab fragments (1:50, Roche) was added and incubated for 3h at room temperature and washed 3 times with PBS.

4.2.14. Capture hybridization analysis of RNA targets (CHART)

The genomic binding sites of *mHotairm* lnc RNA were identified by using the CHART technique developed by Simon *et al*., (2011) and was modified as described here. NIH 3T3 cells (10^7) were collected and cross-linked in 10 ml PBS with 1% formaldehyde for 10 min at room temperature. Cells were washed 3 times with PBS and dounce homogenized in 4 ml sucrose buffer (0.3M sucrose, 1% Triton X-100, 10mM HEPES pH 7.5, 100 mM KOAc, 0.1 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine, 1× protease inhibitor, 1 mM DTT, 10 unit/ml RNase

inhibitor). The homogenized sample was added to the top of 4 ml glycerol buffer (25% glycerol, 10 mM HEPES pH 7.5, 100mM KOAc, 1 mM EDTA, 0.1 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine, 1× Roche protease inhibitor tablet, 1 mM DTT, 10 unit/ml RNase inhibitor) and centrifuged at 1000g for 15 min at 4°C. The pellet (nuclei) was collected and cross-linked in 10 ml PBS with 3% formaldehyde for 30 min at room temperature. Nuclei pellet was then washed with PBS for 3 times and resuspended in 1 ml WD100 solution (100mM NaCl, 10Mm Hepes PH 7.5, 2Mm EDTA, 1Mm EGTA, 0.2% SDS and 0.1% N-Lauroylsarcosine). The resuspended nuclei pellet was sonicated using Branson digital sonifier 250 (Branson) with 10% input pulse, 3 cycles with 2 s pulse "on" and 30 s pulse "off" to shear DNA to 2-3 kb fragments. Sheared sample was centrifuged at 16000 g for 10 min at room temperature and 50 µl supernatant was collected as 5% input. The rest of supernatant was adjusted to hybridization conditions (20 mM HEPES, pH 7.5, 817 mM NaCl, 1.9 M urea, 0.4% SDS, 5.7 mM EDTA, 0.3 mM EGTA, 0.03% sodium deoxycholate, 5× Denhardt's solution) and pre-cleared with ultralink-streptavidin resin (Pierce). Three biotin labeled oligo probes against *mHotairm1* were designed (Fig. 4.4, Table 4.3) and were added (16 nM each) and hybridized with 55 °C for 20 min, 37 °C for 10 min, 45 °C for 60 min and followed by 37 °C for 30 min. The bound material was captured using streptavidin beads overnight at room temperature. The product was rinsed five times with WB250 (250 mM NaCl, 10 mM Hepes, pH 7.5, 2 mM EDTA, 1 mM EGTA, 0.2% SDS, 0.1% N-lauroylsarcosine). For protein sample detection, beads were suspended in RIPA buffer and boiled with 4x loading buffer for SDS-PAGE and western blot analysis as described in 4.2.6. For DNA sample detection, 100 µl reverse buffer (1mg/ml proteinase K, 0.5% SDS and 100Mm Tris pH 7.4) was added to the beads and incubated at 55°C for 1 h followed by 65 °C for 30 min. DNA was then isolated using GeneJET PCR Purification kit (ThermoFisher Scientific) and *Hoxa1*, *Hoxa2*,

Hoxa3, *Hoxa5* and *Hoxa13* promoters were identified by PCR using primers listed in Table 4.3 (Simon *et al*., 2011).

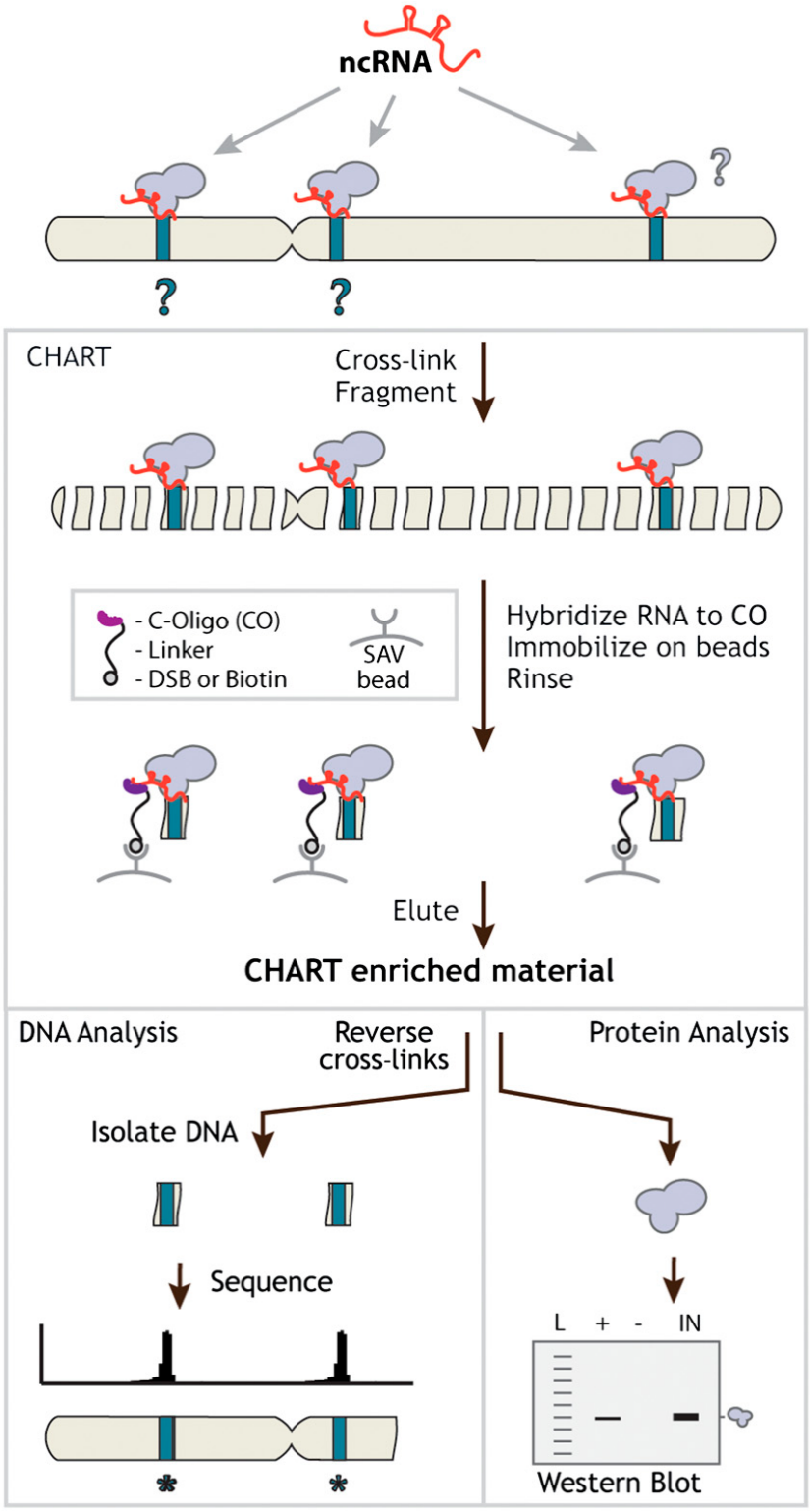


Figure 4.5. Capture hybridization analysis of RNA targets (CHART). Chromatin was crosslinked and sheared. Biotin labelled complementary oligonucleotides were used to purify the targeted RNA together with its DNA and protein targets using streptavidin beads. Reversibly cross-linked CHART-enriched material can then go through different analysis, e.g. DNA analysis (left) and protein analysis (right) (taken from Simon *et al*., 2011 with permission).

4.2.15. GST fusion protein pull down experiments

4.2.15.1 GST fusion protein expression vector construction

WDR5 and MLL1 (amino acid 3810-3963) were amplified from NIH 3T3 cDNA samples using PCR primers containing EcoR I (GAATTC) and BamH1 (GGATCC) restriction enzyme digestion sites (Table 4.3, Mouse MLL13810 start, Mouse MLL1 3963 stop, Mouse WDR5 start, Mouse WDR5 stop). PCR amplification system (a total of 50 µl) were prepared with 50 ng cDNA, 10×EconoTaq buffer, 2.5 units of EconoTaq DNA polymerase (Invitrogen), 50 µM of each dNTP and 500 nM of each forward and reverse primer (Table 4.3). The PCR cycle started with one cycle at 95°C for 3 min, followed by 30 cycles of 95°C for 30 sec, 53°C for 40 sec, 72°C for 1 min and a final extension step at 72°C for 10 min. The PCR products were run on a 1.0% agarose gel and purified from the gel using QIAquick Gel Extraction kit (Qiagen). The DNA samples were then cloned into pGEM T-easy vector (Promega) and sequenced with an ABI PRISM™ system at the National Research Council, Plant Biotechnology Institute, Saskatoon, SK. The correct clone sequences were then digested from T-easy vector using FastDigest™ *Eco*RI and FastDigest™ *Bam*HI restriction enzymes (Thermo Scientific™) at 37°C for 20 min. The digested DNA fragments were run on a 1% agarose gel and target bands were excised from the gel and purified using QIAquick Gel Extraction kit (Qiagen). The purified DNA samples were then cloned into pGEX-6p-1 vector (Fig. 4.6. GE Healthcare). This started with the incubation of 100 ng pGEX-6p-1 vector with 50 ng of purified DNA fragment (WDR5 or MLL1) at 65°C for 5 min. Vector/DNA fragment mix were then placed on ice and a ligation (10 µl) was

set up by adding 5 µl of 2x Ligation buffer and 0.5 µl of T4 DNA ligase (Promega) to the mix. Ligation was carried at 16°C overnight followed by 5 min incubation at 70°C to inactivate T4 DNA ligase. Ligation product was then transformed into DH5α™ Competent Cells (Subcloning Efficiency™, Invitrogen) as described in 4.2.2.3. Single clone was picked and incubated in 4ml LB medium with 1% ampicillin overnight at 37°C. Plasmid was then extracted from bacteria culture and target fragment insertion was tested by digesting the plasmid DNA with FastDigest™ *Eco*RI and FastDigest™ *Bam*HI restriction enzymes (Thermo Scientific™) at 37°C for 20 min. The digested DNA fragments were run on a 1% agarose gel. DH5α™ cells with correct WDR5 and MLL1 expression vectors were stored at -80°C in LB medium with 25% glycerol.

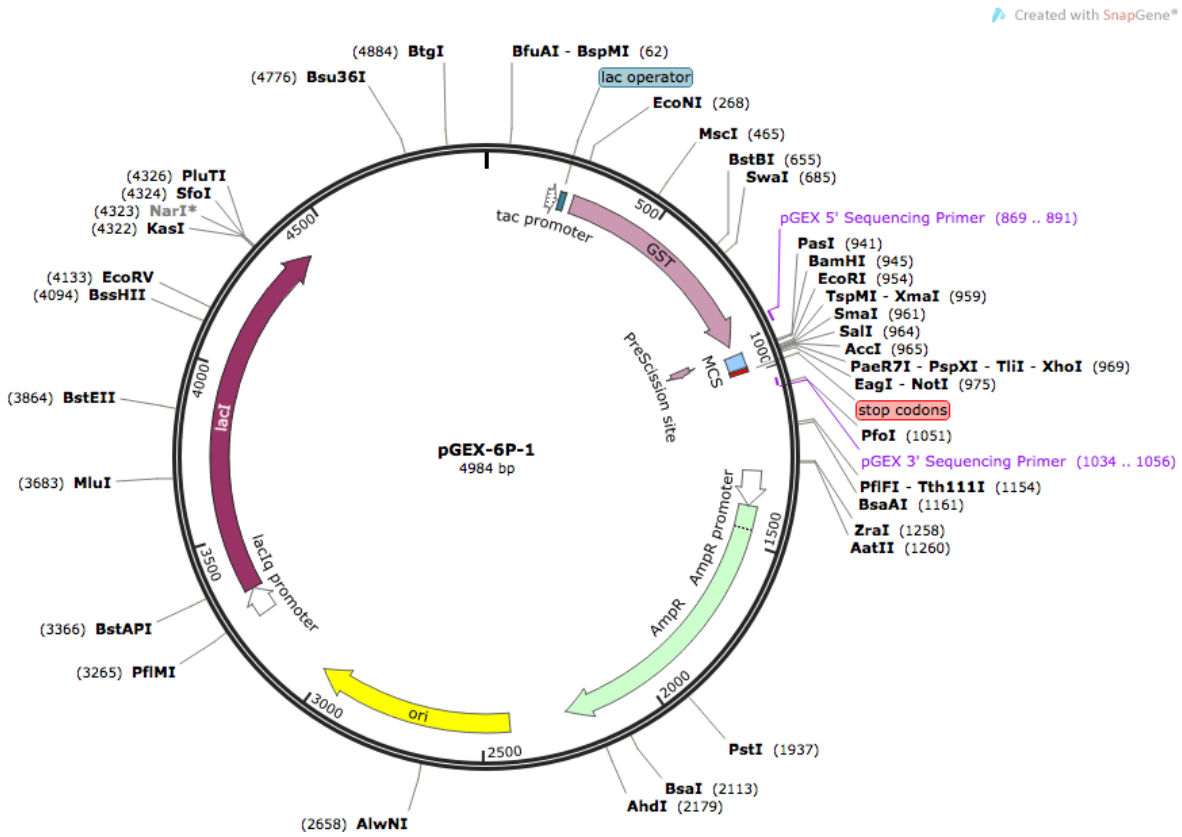


Figure 4.6. pGEX-6p-1 vector map (GE Healthcare). WDR5 and MLL1 (amino acid 3810-3963) were cloned into pGEX-6p-1 vector using *Eco*RI and *Bam*HI restriction enzymes to link with GST (Figure taken from www.snapgene.com/resources)

4.2.15.2 GST fusion protein expression and purification

The bacterial DH5α™ cells with WDR5 and MLL1 expression vectors were cultured in 4 ml LB medium with 1% ampicillin overnight at 37°C. The cells were then subcultured (1:100 v/v) into LB medium with 1% ampicillin for 2 h at 37°C. Protein expression was induced following treatment with IPTG (1mM) for 2h at 26°C.

DH5α™ cells were collected by centrifuging at 3000rpm for 3 min and vortexed with 1ml lysis buffer (1mM DTT, 0.5 mg/ml lysozyme, 1% triton X-100 and 1x proteinase inhibitor in RIPA buffer) and incubated on ice for 20 min. The samples were then sonicated using Branson digital sonifier 250 (Branson) (10% input pulse, 5 cycles with 2 s pulse "on" and 30 s pulse "off"). Following sonication, samples were centrifuged at 12000 rpm for 10 min at 4°C and supernatant was collected. Glutathione agarose beads were washed with RIPA buffer. A 100 µl of glutathione agarose beads (Pierce™, ThermoFisher Scientific) were added into each supernatant collected from the previous step and incubated for 2 h at room temperature to purify GST fused protein. The beads were collected by centrifuging at 3000rpm for 1 min and washed for three times with RIPA buffer. The purification of GST fused protein was tested by running the purified beads on an SDS-PAGE gel, followed by coomassie blue staining.

4.2.15.3 GST fusion protein pull down

For protein sample pull down, the purified glutathione agarose beads with protein were added to NIH 3T3 cell lysates (in RIPA buffer) and incubated at 4°C overnight. Pull down samples were collected by centrifuging at 3000rpm for 2 min and washed three times with RIPA buffer. Protein samples pulled down with GST fused MLL1 (amino acid 3810-3963) were tested

with WDR5 antibody following western blot analysis as described in 4.2.6. GST was used as a negative control.

For RNA sample pull down, NIH 3T3 cells were lysised in cell lysis buffer with 100mM KCl, 5mM $MgCl_2$, 10mM Hepes (PH 7.0), 0.5% Nonidet P-40, 1mM DTT, 2mM Vanadyl ribonucleoside complexes solution, 25 μg/ml protease inhibitor and 100 unit/ml RNase inhibitor. Lysates were votexed and centrifuged at 12000rpm at 4°C for 10 min. Purified glutathione agarose beads with WDR5/MLL1 were added into the supernatant and incubated at 4°C for 3 h. Beads were then collected by centrifuging at 3000rpm for 2 min and washed 3 times with cell lysis buffer. RNA samples were recovered using RNeasy Protect Mini Kit (Qiagen) and converted to cDNA as described in 4.2.4. The existence of *mHotairm1* was tested by PCR using mush 138 forward and reverse primers as listed in Table 4.3.

4.2.16 Chromatin Immunoprecipitation (ChIP)

NIH 3T3 cells transfected with either *hotairm1* siRNA or control siRNA were cross linked with 1% formaldehyde in cell culture media for 10 min at room temperature. Glycine was added to a final concentration of 125 mM and incubated with shaking for 5 min at room temperature. Cells were washed two times with 10 ml cold PBS. Cells were collected and lysed with Lysis Buffer (50Mm HEPES-KOH pH7.5, 140Mm NaCl, 1Mm EDTA pH8, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS and protease inhibitor cocktail). Cell lysates were sonicated using Branson digital sonifier 250 (Branson) with 20% input pulse, 3 cycles of 5s pulse "on" and 30s pulse "off" to shear DNA to lengths between 200 and 1000 bp. Sheared DNA samples were centrifuged for 30 s at 8,000 g at 4°C and supernatant was collected. Input represents 5% of supernatant.

Samples were then incubated with 10 µg anti-Histone H3 (tri methyl K4) antibody (Abcam), 10 µg H3K27trime antibody (Abcam), 10 µg Histone3 antibody (Abcam) or 10µg normal rabbit IgG (Santa cruz) respectively, at 4°C overnight. Antibody/histone complex were collected with 20 µl protein A agarose slurry (Santa Cruz). Beads were washed three times with 1 ml wash buffer (0.1% SDS, 1%Triton X-100, 2mM EDTA pH8, 150mM NaCl and 20mM Tris-HCl pH8). Histone complex from the antibody was eluted with 100 µl elution buffer (1% SDS and 100 mM $NaHCO_3$) and histone-DNA crosslinks were reversed by heating at 65°C in 5M NaCl for 4 h. Resulting DNA was then recovered by incubating with proteinase K and extracted by phenol/chloroform. Samples were resuspended in 20 µl of MiliQ distil water, and 1:50 was used for qPCR with *Hoxa1*, *Hoxa2* promoter primers listed in Table 4.3. qPCR was performed in a final volume of 25 µl containing 50 ng of ChIP DNA or 50 ng of input DNA as template, 400 ng of each primer, and 12.5 µl of SYBR green PCR master mix (Applied Biosystems). The amplification consisted of 10 min at 95°C followed by 35 cycles of 95°C for 30 s, 60°C for 1 min, and 72°C for 30 s in 7300 ABI detection system. The fold enrichment of each protein at the *Hoxa1* and *Hoxa2* gene promoter was determined by the 2(-ΔΔCt) method described in the Applied Biosystems User Bulletin (ABI PRISM 7700 Sequence Detection System, Applied Biosystems, 2001) and are shown as the fold increase relative to input, N=3.

4.2.17 Separation of nucleus and cytoplasm

NIH 3T3 cells were collected and washed in cold PBS. The cells were homogenized in 0.5 ml sucrose buffer (0.3M sucrose, 1% Triton X-100, 10mM Hepes pH 7.5, 100 mM KOAc, 0.1 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine, 1× Roche protease inhibitor tablet, 1 mM DTT, 10 unit/ml RNase inhibitor). The homogenized sample was added to the top of 0.5 ml glycerol

buffer (25% glycerol, 10 mM Hepes pH 7.5, 100mM KOAc, 1 mM EDTA, 0.1 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine, 1× protease inhibitor, 1 mM DTT, 10 unit/ml RNase inhibitor) and centrifuged at 1000g for 15 min at 4°C. The supernatant was collected as cytoplasm and the pellet was collected as nucleus (Simon *et al*., 2011).

4.2.18 Immunoprecipitation (IP)

NIH 3T3 cells were lysed in RIPA buffer with protease inhibitor cocktail (Thermo Fisher Scientific) and sumoylation/ubiquitilation protector N-ethylmaleimide (NEM). Cell samples were collected and vortexed and kept on ice for 5 min. Agarose protein A beads were pre-cleaned with RIPA buffer. Cell lysate was centrifuged for 15 min at 12000 g at 4°C and supernatant collected. For WDR5 antibody IP, 10µg WDR5 antibody (ab22512, rabbit IgG, Abcam) was added to the cell lysate to precipitate WDR5 at 4°C over night. Normal rabbit IgG antibody (10µg, Santa Cruz) was used as negative control. Pre-cleaned agarose protein A beads were then added and samples were incubated at 4°C on shaker for 1h. Samples were centrifuged at 8000 g for 30 s and the precipitates were washed three times with RIPA buffer. The collected samples were denatured at 100 °C for 10 min for subsequent SDS-PAGE (10%) separation. Protein samples were transferred and immobilized onto a PVDF transfer membrane. Membrane was blocked in 3% BSA (in PBS) at 4 °C overnight. SUMO-1 antibody (1:500, Developmental Studies Hybridoma Bank) was used to detect sumoylated WDR5. For SUMO1 antibody IP, 10µg SUMO1 antibody (mouse IgG, Developmental Studies Hybridoma Bank) was used in IP and normal mouse IgG (Santa Cruz) was used as negative control. Western blot analysis was carried out using WDR5 antibody (WDR5 A-6, mouse IgM, Santa Cruz).

4.2.19 Statistical data analysis

All statistical analyses were performed using the GraphPad 5.0 software (GraphPad Prism). Statistical analysis on qRT-PCR of miRNAs and gene expressions were evaluated using one way analysis of variance (ANOVA) followed by Bonferroni post-tests. Student's T-Tests was used in the analysis of ChIP experiment and luciferase assay for the comparison between experimental groups and control groups. A significant p-value of less than or equal to 0.05 was used to denote significant difference

Table 4.3. Sequences of primers and probes used in this study.

Name	Sequence
Hoxa2 gene (Taqman)	Forward: 5' CTGGATGAAGGAGAAGAAGGC Reverse: 5' CGGTTCTGAAACCACACTTTC
GAPDH gene (Taqman)	Forward: 5' ACCACAGTCCATGCCATCAC Reverse: 5' TCCACCACCCTGTTGCTGTA
β-actin gene (SYBR Green)	Forward: 5' AGAGGGAAATCGTGCGTGAC Reverse: 5' CAATAGTGATGACCTGGCCGT
Methylation specific PCR primers	Msp- CpG1-MF:5'TTTTCGATAGTTTTAAATAATGCGC Msp- CpG1-MR:5'TAAATAACAATACCCCGAAAATACG Msp-CpG1-UF:5' TTTGATAGTTTTAAATAATGTGTGG Msp- CpG1-UR:5'AATAACAATACCCCAAAAATACATA Msp- CpG2-MF:5' TAGTTATTTTTGAGAAGTTGATGGC Msp- CpG2-MR:5' TATCCAACCCGAAACCTACG Msp- CpG2-UF:5'GTTATTTTTGAGAAGTTGATGGTGA Msp- CpG2-UR:5' TTAATATCCAACCCAAAACCTACAC Msp- CpG3-MF:5' TTTTTGGAAGTCGGAAATATTTATC Msp- CpG3-MR:5'AATCCCCGTAACCTAAACGAC Msp-CpG3-UF: 5' TTTTGGAAGTTGGAAATATTTATTGT Msp-CpG3-UR:5' CAAATCCCCATAACCTAAACAAC
Bisulfite specific PCR primers	Bsp-CpG3-F: 5' GTTAGGTTGAGGTGTTTAAAT Bsp-CpG3-R: 5' ACCACTATCTATTTAATTAATCC
mHotairm1 primers	mush138-F: 5'CCCACCCAGCCCAGAAAG mush138-R: 5'GTTTCAAACATCTACGTTCC mush 147-F: 5'TGACTTGGAGCACTGGGA mush 147-R: 5'CTCTTGCCAGTTCAGCTTTCT mush 245-F: 5' GAAAGGAAGAGTTGGAACGTAGA mush 245-R: 5' TGAGACTCAGGCCATAGAGTTA

Hoxa2 3' UTR mutation primers	669b1F: 5' CACACGACAAAACGCCTTTGACC 669b1R: 5' GGTCAAAGGCGTTTTGTCGTGTG 669b2F: 5' CTATGTGATTTTCCTGAAAAACACGACAGGAGGCCTGC 669b2R: 5' GCAGGCCTCCTGTCGTGTTTTTCAGGAAAATCACATAG
GST fusion protein Mouse MLL1 primers	3810start: 5'**GGATCC**ATGCCCATGAGATTCCGGCACTTG 3963stop: 5'**GAATTC**TTAGTTCAGGAACTTGCGGCA
GST fusion protein Mouse WDR5 primers	start: 5' **GGATCC**TTCAGAGCCATGGCCACAGAGG stop: 5' **GAATTC**TTAGCAGTCACTCTTCCACAGT
Promoter primers *Hoxa1*	Forward: 5' TTTGGTCCCAGTGCTCCAAG Reverse: 5' ATGTACAGTGCGCAAGAGGC
Hoxa2	Forward: 5' GTTGTCTTTTGAATCCTTCAGCC Reverse: 5' GGCACTCTGGGTTTCATACC
Hoxa3	Forward: 5' CTCAAAAGGGCAAAGGGTCG Reverse: 5' GGTTGGTCCTCTTCCCTGAAC
Hoxa5	Forward: 5' TCCTAATGGAACTGCGAGGG Reverse: 5' CTAATGGGGGAGTTGGGTGG
Hoxa13	Forward: 5' AACCAACAGGAAACAAACGCC Reverse: 5' GGTCGTGGGGACAAGTCAG
CHART probes	Probe 1: Biotin-5' GGCGCGCGCAGAAGGCAGTCTCG Probe 2: Biotin-5' TCAAACATCTACGTTCCAACTCT Probe 3: Biotin-5' AGACTCAGGCCATAGAGTTACAT Probe C: Biotin-5' AGAGTTGGAACGTAGATGTTTGA (Antisense to Probe 2)
mHotairm1 siRNA	5' GGAAGAGUUGGAACGUAGAUGUUTG
In *situ* Probes *mHotair1*	5' DigN/ACTCCGTTATTGACCTAGAAACTAGCAGCTTGGTAAGGGAAC

CHAPTER 5

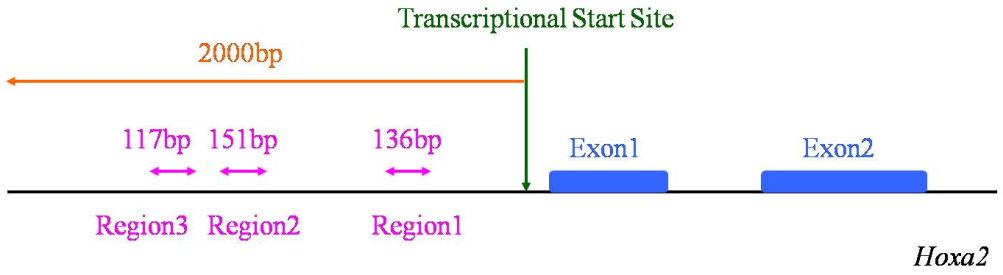
5. Results and analysis of data

5.1 Methylation status of *Hoxa2* promoter region in developing mouse palate, NIH 3T3 cells and EG7 cells.

DNA methylation was the first epigenetic mechanism that I investigated in the regulation of *Hoxa2* gene expression in mouse palate development. Analysis of the *Hoxa2* promoter revealed the presence of three CpG islands (Fig. 5.1). The three sets of methylation specific primers used for the analysis of the CpG islands are listed in the materials and methods section (Table 4.3). Methylation specific PCR (MSP) carried out for all three CpG rich regions showed that all three CpG regions in the *Hoxa2* promoter were unmethylated in developing palate as well as in NIH 3T3 cells, while these regions remained methylated in EG7 cells, (Fig. 5.2). A further bisulfite specific PCR (BSP) amplification and sequencing was carried out to examine the methylation status of each CpG site. A region close to the transcription start site (Region 1, Fig. 5.1) with 14 CpG sites was selected for further analysis. I found that all 14 CpG sites were unmethylated in both NIH 3T3 cells and in developing palate samples (Fig. 5.3A, Appendix 1-5); however, in the same CpG rich region 1 of the *Hoxa2* promoter in EG7 cells, 10 of the 14 sites were methylated (Fig. 5.3B, Appendix 6).

Since *Hoxa2* is expressed in both the developing mouse palate (Smith *et al*., 2009) and NIH 3T3 cells but not in EG7 cells (X. Wang, PhD Thesis, 2013), my results indicate that an unmethylated promoter of *Hoxa2* may be required for its expression. However, since *Hoxa2* expression changes temporally at different stages of mouse palate development (Smith *et al*., 2009), and that the *Hoxa2* promoter remains unmethylated throughout mouse palate

development, DNA methylation does not appear to be a regulator for *Hoxa2* expression during mouse palatogenesis, therefore, I focused on other epigenetic regulators, such as miRNAs and lncRNAs.



Region 1 (-489 -- -624bp 136bp)
CTGCAGGCGTCAGGCTGAGGTGCTTAAATGATTTGTGAGGTGCGAGGCGTCTTCCCG
ACAGTCCCAAACAATGCGCGGAGTGTGCGGGGGAGGCAGAGGGCAGCCACTGGCG
GGACGGCAGCAGGGCTCACACGCA

Region 2 (-1019 -- -1169bp 151bp)
GTTGATGGCGAAGGAAGATCAGCAGGCTTCGAGAGCCGCCTCTCGTTTTCCGCTGCC
CGCAGGAGTCAGATGCAAAGCTGCCGTGGAAAAGCCGGCTAACAATGGGCCAGGG
GCGCAGGTCCCGGGCTGGACATTAAGAGGAGGCGAGAGG

Region 3 (-1244 -- -1360bp 117bp)
AAGTCGGAAATACTCACCGCACCCGAGCCCTACGGGTATGAAACCCAGAGTGCCAG
GAGCCGCACGAGCCTGCTCGGGACGGCATTGTTTTGGCTGGCCGCCGCCCAGGCTAC
GGGG

Figure 5.1. Analysis of CpG rich regions in *Hoxa2* gene promoter. All the predicted CpG sites in the region 2000 bp upstream of the transcriptional start site were separated into three methylation specific regions: region 1 from -489 to -624, region 2 from -1019 to -1169, and region 3 from -1244 to -1360 (transcriptional start site begins at 0). CpG sites are highlighted with light grey. Predicted methylated cytosines of CpG dinucleotides are shown as red letters (predicted with Methylator, Bhasin *et al*., 2005, http://bio.dfci.harvard.edu/Methylator/).

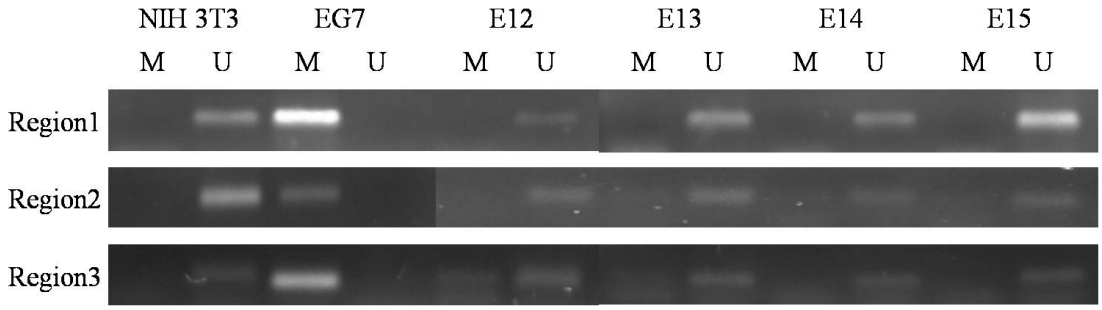


Figure 5.2. MSP amplification for three CpG islands in *Hoxa2* promoter. MSP was done for DNA samples collected form NIH 3T3 cells, EG7 cells and developing mouse palatal tissue (E12-E15). M = PCR fragment amplified with designed methylated primers. U = PCR fragment amplified with designed unmethylated primers. The presence of a PCR band in M indicates the region is methylated and PCR band in U indicates the region is unmethylated.

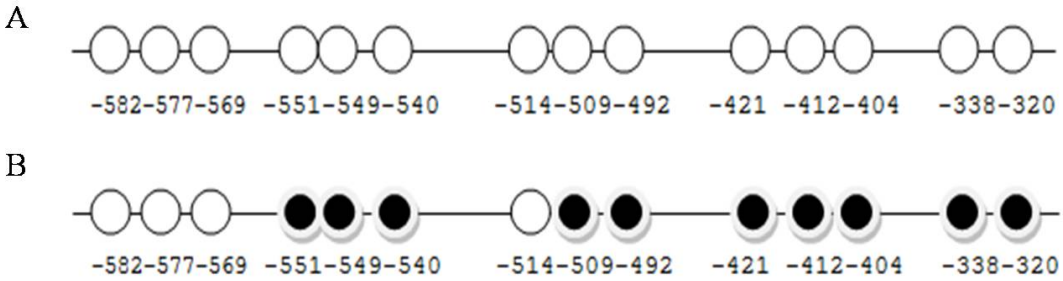


Figure 5.3. BSP and sequencing results for the CpG rich region 1. (A). Schematic of CpG methylation status in NIH 3T3 cells and in E12 to E15 palatal samples. All 14 CpG sites remained unmethylated. (B). Schematic of CpG methylation status in EG7 cells. 10 out of 14 CpG sites were methylated in EG7 cells. Each circle represents a CpG site. The number under each circle represents the position of each CpG site on mouse *Hoxa2* promoter (transcriptional start site begins at 0). Blank circles: unmethylated CpG site; Black circles: methylated CpG site.

5.2 Regulation of *Hoxa2* gene expression by miRNAs

5.2.1 miRNA binding sites prediction and sequence analysis.

MicroRNAs play an important role in the regulation of *Hox* gene expression in several biological and pathological processes, mainly via binding to the 3' UTR of *Hox* genes (reviewed in section 1.2.3). To study the regulation of *Hoxa2* gene expression by miRNAs, I first used an online software (http://www.microrna.org) to predict miRNAs that bind to mouse *Hoxa2* 3' UTR. Six miRNAs were predicted to bind to mouse *Hoxa2* 3' UTR as shown in Figure 5.4. The stem-loop sequence and mature sequence of all six miRNAs are shown in Table 5.1. Detailed predicted binding sequences and structures between miRNAs and mouse *Hoxa2* 3' UTR are shown in Figures 5.5-5.10 (A, B) for each individual miRNA. I also analyzed the sequence similarity for each miRNA binding site among several species (Figs. 5.5–5.10C). Most predicted miRNA binding sites are evolutionally conserved except the binding site 2 of miR-669b (Fig. 5.5C)

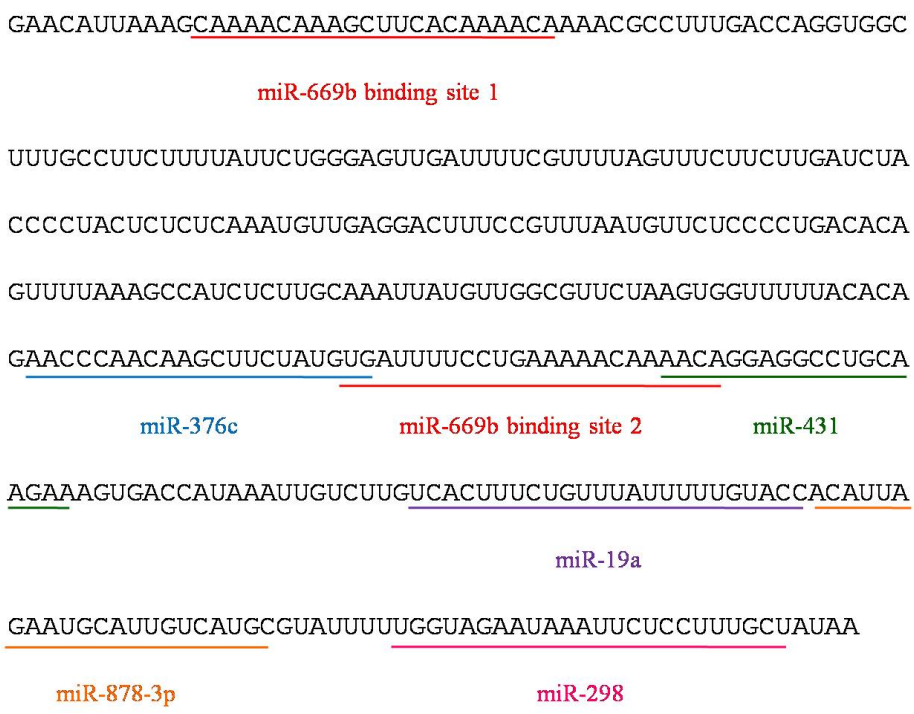
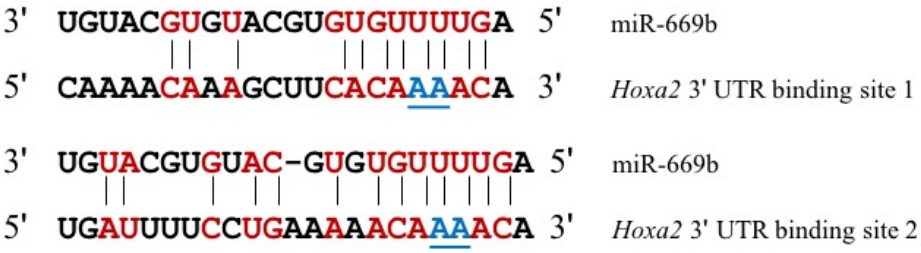


Figure 5.4. Predicted binding sites of miRNA on mouse *Hoxa2* 3' UTR (prediction from http://www.microrna.org). Full length mouse *Hoxa2* 3' UTR sequence is shown (5' to 3'). Colour lines beneath the nucleotide sequence represents binding sites of respective miRNAs.

Table 5.1. Mouse miRNA stem-loop sequence and mature sequence. The sequence data is from miRBase (http://www.mirbase.org/).

	Stem-loop sequence 5'-3'	Mature sequence 5'-3'
miR-669b	AUGAAUGUAUGUGCAUGUGUA UAUAGUUUUGUGUGCAUGUGC AUGUGUGUCUAUUAAUGUACA UAUACAUACACACAAACAUAU ACACGCAUGCGCA	AGUUUUGUGUGCAUGUGCAU GU
miR-431	CGUCCUGCGAGGUGUCUUGCA GGCCGUCAUGCAGGCCACACUG ACGGUAACGUUGCAGGUCGUC UUGCAGGGCUUCUCGCAAGAC GACAUC	UGUCUUGCAGGCCGUCAUGCA
miR-298	CCAGGCCUUUGGCAGAGGAGG GCUGUUCUUCCCUUGAGUUUU AUGACUGGGAGGAACUAGCCU UCUCUCAGCUUAGGAGUGG	GGCAGAGGAGGGCUGUUCUU CCC
miR-376c	UUUGGUAUUUAAAAGGUGGAU AUUCCUUCUAUGUUUAUGCUU UUUGUGAUUAAACAUAGAGGA AAUUUCACGUUUUCAGUGUCA AA	AACAUAGAGGAAAUUUCACG U
miR-19a	GCAGCCCUCUGUUAGUUUUGC AUAGUUGCACUACAAGAAGAA UGUAGUUGUGCAAAUCUAUGC AAAACUGAUGGUGGCCUGC	UGUGCAAAUCUAUGCAAAAC UGA
miR-878-3p	UGCAAUGCUUUAUCUAGUUGG AUGUCAAGACACGUGAAACUU AAGUGCAUGACACCACACUGG GUAGAGGAGGGCUCA	GCAUGACACCACACUGGGUAG A

A



B

miR-669b binding site 1

miR-669b binding site 2

C

miR-669b	3' UGUACGUGUACGUGUGUUUUGA 5'
Human	5' AUUAAAGCAAAACAAAGCAUCACCAAACAAAAACUCCUU 3'
Mouse	5' AUUAAAGCAAAACAAAGCUUCACAAAACAAAACGCCUU 3'
Rat	5' AUUAAAGCAAAAGAAAGCUUCACAAAACAAAACGCCUU 3'
Chimpanzee	5' AUUAAAGCAAAACAAAGCAUCACCAAACAAAAACUCCUU 3'
Dog	5' AUUAAAGCAAAACAAAGCAUCACAAAACAAAACCCCCUU 3'

Hoxa2 3' UTR miR-669b binding site 1

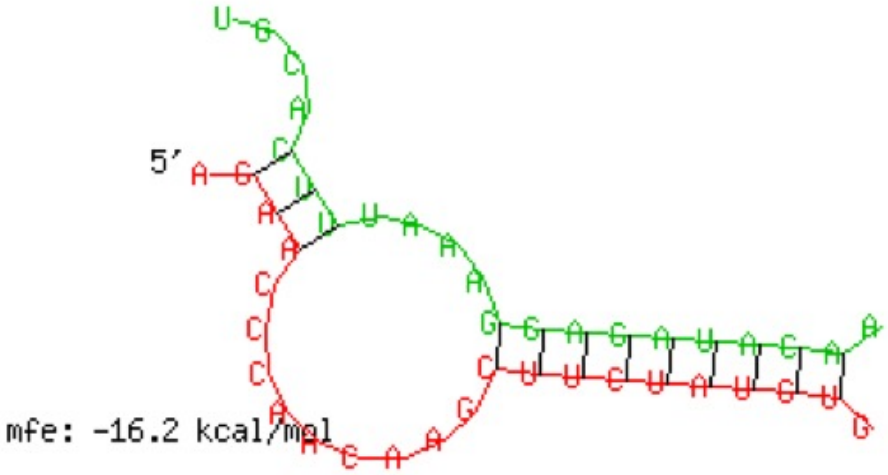
miR-669b	3' UGUACGUGUACGUG-UGUUUUGA 5'
Human	5' CCUUUAUGUGAUU--CCUGAGA-GCA--GUAUGAGGCCU 3'
Mouse	5' CUUCUAUGUGAUUUUCCUGAAAAACAAAACAGGAGGCCU 3'
Rat	5' CUUCUAUGUGAUUUUCCUG-AAAACAAAACACGAGGCCU 3'
Chimpanzee	5' CCUUUAUGUGAUU--CCUGAGA-GCAA--UAUGAGGCCUGCAAG 3'
Dog	5' CCUCUAUGUGAUUU-CCUGAAA-GCAAUAUAUGAGGCCUGC 3'

Hoxa2 3' UTR miR-669b binding site 2

Figure 5.5. Prediction and analysis of miR-669b binding sites on *Hoxa2* 3' UTR. (A) The predicted target seed sequences for miR-669b in mouse *Hoxa2* 3' UTR is marked in red (http://www.microrna.org). The underlined "AA" in blue in both binding sites were mutated to "CG" for luciferase assays. (B) Predicted binding structures of miR-669b to mouse *Hoxa2* 3' UTR (RNAhybrid, Rehmsmeier *et al*., 2004). mfe: minimal free energy. *Hoxa2* 3' UTR sequence is labelled in red and miR-669b is labelled in green. (C) The degree of sequence similarity within the seed targets for miR-669b were shown for *Hoxa2* genes from five different species. Complementary regions between miR-669b and *Hoxa2* 3' UTR were highlighted in red.

A

B



C

miR-376c	3' UGCACUUUAAAGGAGAUACAA 5'
Human	5' UUUUAAACCUAACUCAACAACCCUUUAUGUGAUU--CC 3'
Mouse	5' UUUUACACAGAACCCAACAAGCUUCUAUGUGAUUUUCC 3'
Rat	5' UUUUACACCGAACCCAACAAGCUUCUAUGUGAUUUUCC 3'
Chimpanzee	5' UUUUAAACCUAACUCAACAACCCUUUAUGUGAUU--CC 3'
Dog	5' UUCUAAAUCUAAUCCAACGACCCUCUAUGUGAUUU-CC 3'

Hoxa2 3' UTR miR-376c binding site

Figure 5.6. Prediction and analysis of miR-376c binding site on *Hoxa2* 3' UTR. (A) The predicted target seed sequences for miR-376c in mouse *Hoxa2* 3' UTR is marked in red (http://www.microrna.org). (B) Predicted binding structures of miR-376c to mouse *Hoxa2* 3' UTR (RNAhybrid, Rehmsmeier *et al*., 2004). mfe: minimal free energy. *Hoxa2* 3' UTR sequence is labelled in red and miR-376c is labelled in green. (C) The degree of sequence similarity within the seed targets for miR-376c were shown for *Hoxa2* genes from five different species. Complementary regions between miR-376c and *Hoxa2* 3' UTR were highlighted in red.

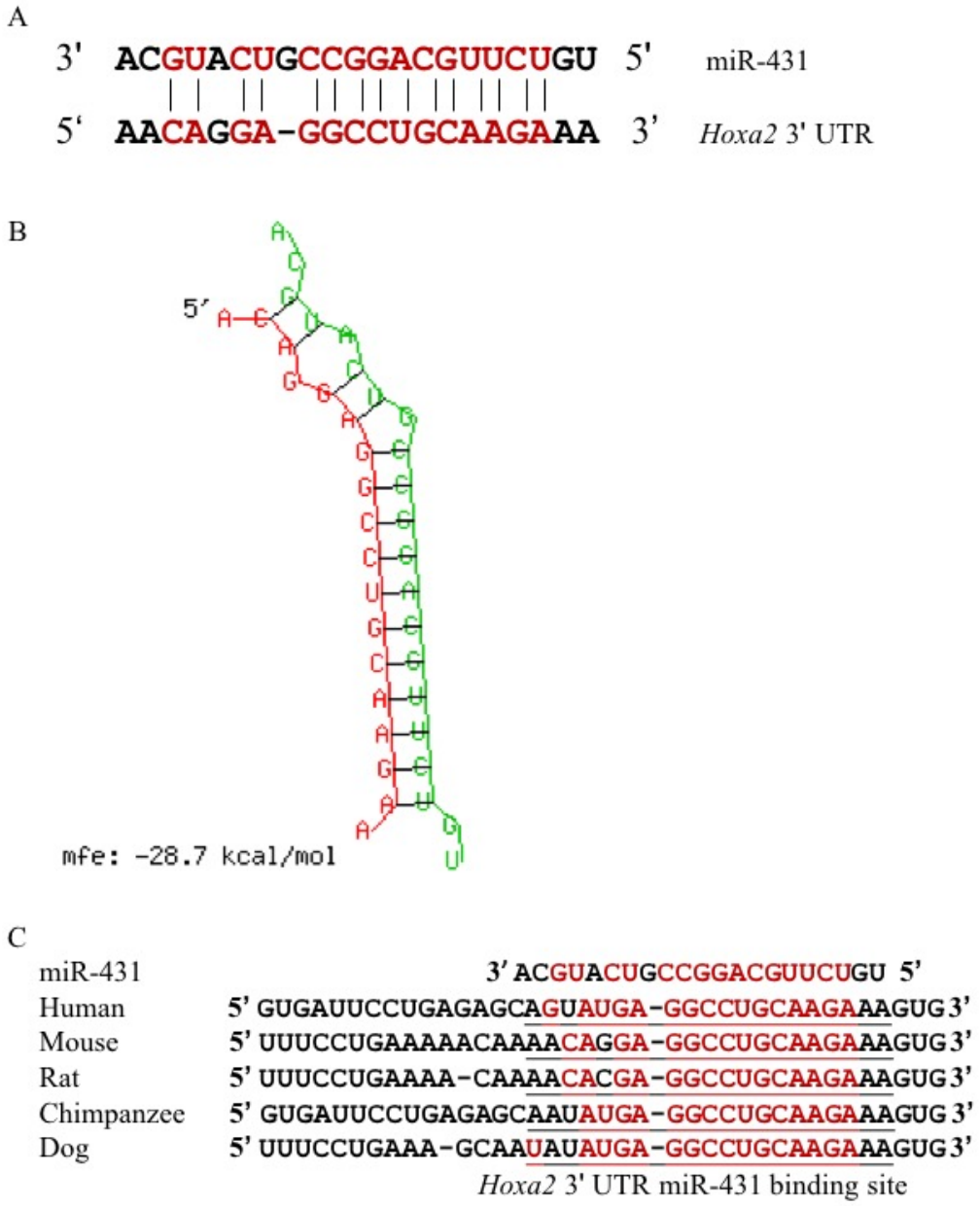


Figure 5.7. Prediction and analysis of miR-431 binding site on *Hoxa2* 3' UTR. (A) The predicted target seed sequences for miR-431 in mouse *Hoxa2* 3' UTR is marked in red (http://www.microrna.org). (B) Predicted binding structures of miR-431 to mouse *Hoxa2* 3' UTR (RNAhybrid, Rehmsmeier *et al*., 2004). mfe: minimal free energy. *Hoxa2* 3' UTR sequence is labelled in red and miR-431 is labelled in green. (C) The degree of sequence similarity within the seed targets for miR-431 were shown for *Hoxa2* genes from five different species. Complementary regions between miR-431 and *Hoxa2* 3' UTR were highlighted in red.

A

B

C

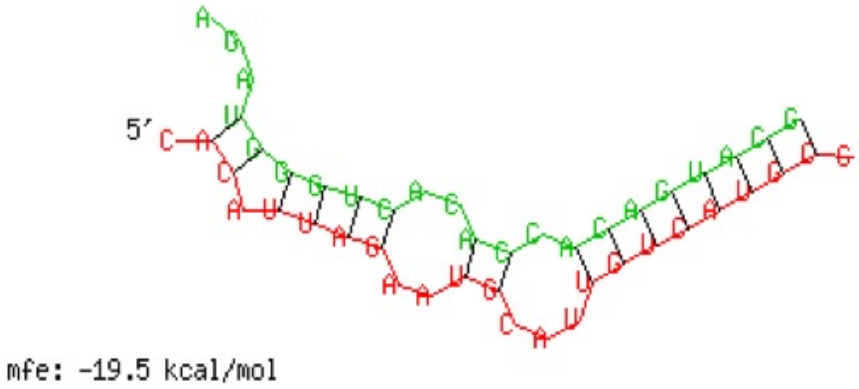
miR-19a	3' AGUCAAA-ACGUAUCUAAACGUGU 5'
Human	5' UUGUAUCUUCACUUUCUUUUUAUUUUUGUAUUACAUUGGGAUGC 3'
Mouse	5' UUGUCUUGUCACUUUCUGUUUAUUUUUGUACCACAUUAGAAUGC 3'
Rat	----------
Chimpanzee	5' UUGUAUCUUCACUUUC---UUAUUUUUGUAUUACAUUGGGAUGC 3'
Dog	5' UUGUAUCUUCACUUUCUUUUUAUUUUUGUAUUACAUUAGGAUGC 3'

Hoxa2 3' UTR miR-19a binding site

Figure 5.8. Prediction and analysis of miR-19a binding site on *Hoxa2* 3' UTR. (A) The predicted target seed sequences for miR-19a in mouse *Hoxa2* 3' UTR is marked in red (http://www.microrna.org). (B) Predicted binding structures of miR-19a to mouse *Hoxa2* 3' UTR (RNAhybrid, Rehmsmeier *et al*., 2004). mfe: minimal free energy. *Hoxa2* 3' UTR sequence is labelled in red and miR-19a is labelled in green. (C) The degree of sequence similarity within the seed targets for miR-19a were shown for *Hoxa2* genes from five different species. Complementary regions between miR-19a and *Hoxa2* 3' UTR were highlighted in red.

A

B



C

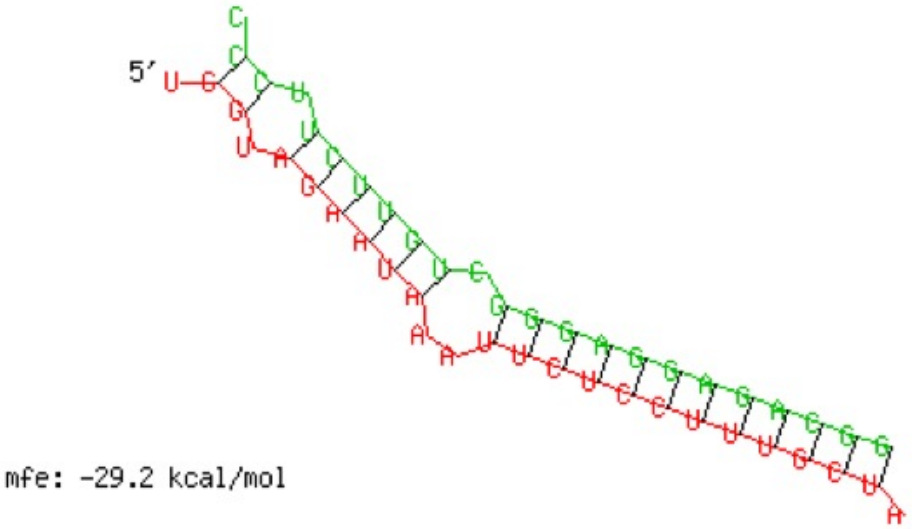
miR-878-3p	3' AGAUGGGUCACACCACUGUACG 5'
Human	5' UUUUAUUUUUGUAUUACAUUGGGAUGCAUUGUCAUGCAUAUU 3'
Mouse	5' GUUUAUUUUUGUACCACAUUAGAAUGCAUUGUCAUGCGUAUU 3'
Rat	----------
Chimpanzee	5' UUUUAUUUUUGUAUUACAUUGGGAUGCAUUGUCAUGCAUAUU 3'
Dog	5' UUUUAUUUUUGUAUUACAUUAGGAUGCAUUGUCAUGAGUAUU 3'

Hoxa2 3' UTR miR-878-3p binding site

Figure 5.9. Prediction and analysis of miR-878-3p binding site on *Hoxa2* 3' UTR. (A) The predicted target seed sequences for miR-878-3p in mouse *Hoxa2* 3' UTR is marked in red (http://www.microrna.org). (B) Predicted binding structures of miR-878-3p to mouse *Hoxa2* 3' UTR (RNAhybrid, Rehmsmeier *et al*., 2004). mfe: minimal free energy. *Hoxa2* 3' UTR sequence is labelled in red and miR-878-3p is labelled in green. (C) The degree of sequence similarity within the seed targets for miR-878-3p were shown for *Hoxa2* genes from five different species. Complementary regions between miR-878-3p and *Hoxa2* 3' UTR were highlighted in red.

A

B



C

miR-298	3' CCCUUCUUG-UCGGGAGGAGACGG 5'
Human	5' CAUGCAUAUUUUUUGUAGAAUAAAUUCUCCUUUGCUAUAA 3'
Mouse	5' CAUGCGUAUUUUUGGUAGAAUAAAUUCUCCUUUGCUAUAA 3'
Rat	----------
Chimpanzee	5' CAUGCAUAUUUUUUGUAGAAUAAAUUCUCCUUUGCUAUAA 3'
Dog	5' CAUGAGUAUUUUUGGUAGAAUAAAUUCUCCUUUGCUA--- 3'
	Hoxa2 3' UTR miR-298 binding site

Figure 5.10. Prediction and analysis of miR-298 binding site on *Hoxa2* 3' UTR. (A) The predicted target seed sequences for miR-298 in mouse *Hoxa2* 3' UTR is marked in red (http://www.microrna.org). (B) Predicted binding structures of miR-298 to mouse *Hoxa2* 3' UTR (RNAhybrid, Rehmsmeier *et al*., 2004). mfe: minimal free energy. *Hoxa2* 3' UTR sequence is labelled in red and miR-298 is labelled in green. (C) The degree of sequence similarity within the seed targets for miR-298 were shown for *Hoxa2* genes from five different species. Complementary regions between miR-298 and *Hoxa2* 3' UTR were highlighted in red.

5.2.2 The predicted miRNAs are expressed in NIH 3T3 and EG7 cells.

To study the role of these miRNAs in the regulation of *Hoxa2* expression, I first needed to determine whether these miRNAs were expressed in the NIH 3T3 cells, EG7 cells and in the developing mouse palate. Expressions of all six miRNAs predicted to bind *Hoxa2* 3' UTR were examined in both NIH 3T3 and EG7 cells using specific miRNA primers (Qiagen®) (Table 4.1). miR-298 has the highest expression in NIH 3T3 cells followed by miR-431. Both miR-298 and miR-431 have low expression in EG7 cells. miR-669b, miR-376c and miR-19a all have higher expression in NIH 3T3 cells than in EG7 cells. In contrast, miR-878-3p is expressed higher in EG7 cells than in NIH 3T3 cells (Fig. 5.11). Generally, the miRNA expression levels in EG7 cells are much lower than in NIH 3T3 cells. One explanation would be because EG7 is a cell line with a high degree of DNA methylation (X. Wang, PhD Thesis, 2013), thus transcription activities may be relatively low compared to that in NIH 3T3 cells. This result fits with the hypothesis that the DNA methylation status can impact the expression of miRNAs (Han *et al*., 2007; Bandres *et al*., 2009; Liu *et al*., 2013). miR-431 and miR-298 are strongly expressed in NIH 3T3 cells, indicating that these miRNAs have a less possibility to regulate *Hoxa2* expression based on the fact that *Hoxa2* is also highly expressed in NIH 3T3 cells.

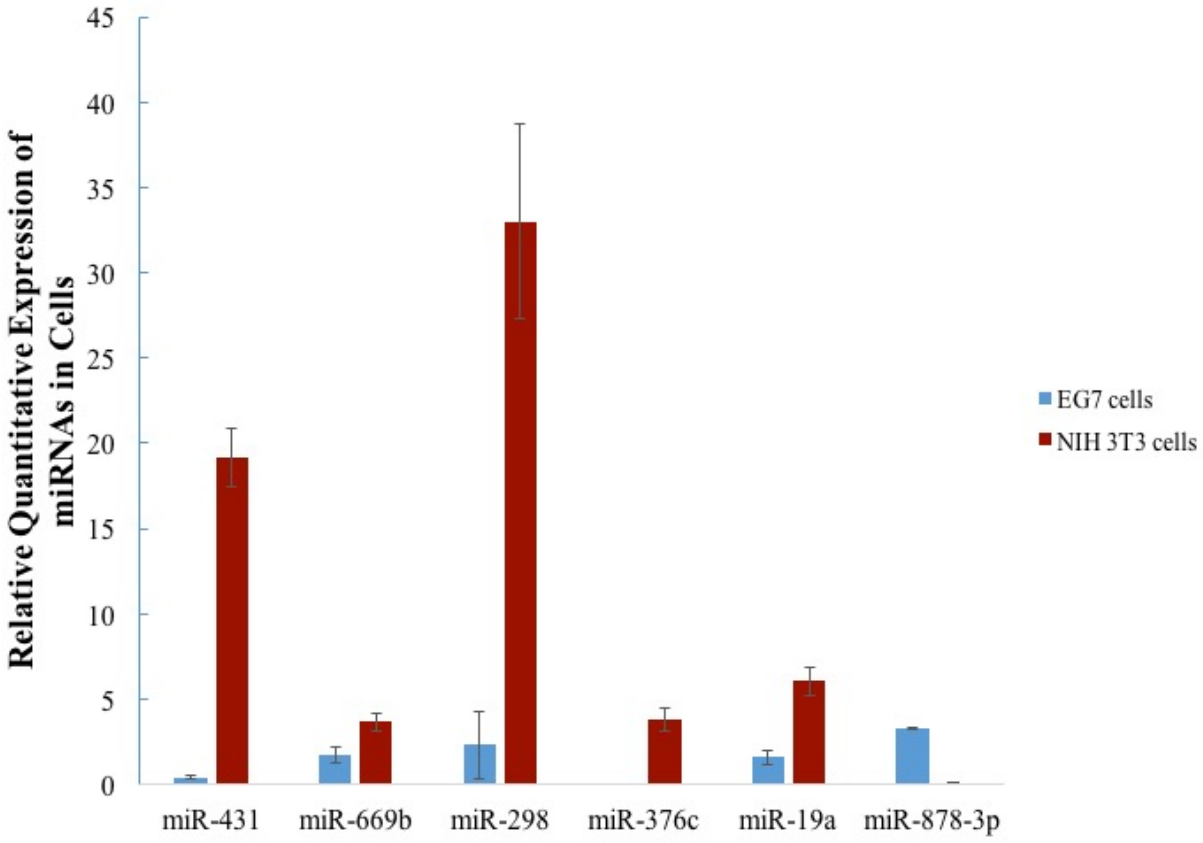


Figure 5.11. miRNA expression in NIH 3T3 and EG7 cell lines. miRNA expression levels in NIH 3T3 and EG7 cells were measured by qRT-PCR as described in section 4.2.5, p76. miRNA expressions were relative to snRNA RNU6B. Relative quantitative expression of miR-431 in EG7 cells were normalized to 1. miR-431 and miR-298 had relatively high expressions in NIH 3T3 cells. Overall, expressions of miRNAs in EG7 cells were relatively low. E - EG7 cells, N - NIH 3T3 cells.

5.2.3 The predicted miRNAs are expressed in the developing mouse palate

All six miRNAs were expressed in the developing mouse palate (Fig.5.12). Interestingly, three of them, miR-669b, miR-376c and miR-431 exhibited gradual increase in expression from stages E12 to E15 during palatal development. These miRNAs had a lower expression at E12, when *Hoxa2* begins its expression in mouse palate (Smith *et al*., 2009). Their expression continued to increase in the developing palate from E13 to E15 (Fig. 5.12) when *Hoxa2* expression gradually declines (Smith *et al*., 2009). The counter and opposite expression levels of these miRNAs to *Hoxa2* gene during mouse palatogenesis indicates that these miRNAs may have a high potential to regulate *Hoxa2* expression during palatal development. For other three miRNAs, miR-19a expression also showed significant differences during some of the stages and may play a role in palate development; however, miR-298 and miR-878-3p did not show significant expression changes during palate development (Fig. 5.12). Taken together with the miRNA expression levels during palate development and miRNA expression in NIH 3T3 and EG7 cell lines, I speculated that miR-669b and miR-376c had a higher possibility to regulate *Hoxa2* expression. Hence, I chose these two miRNAs for further investigation.

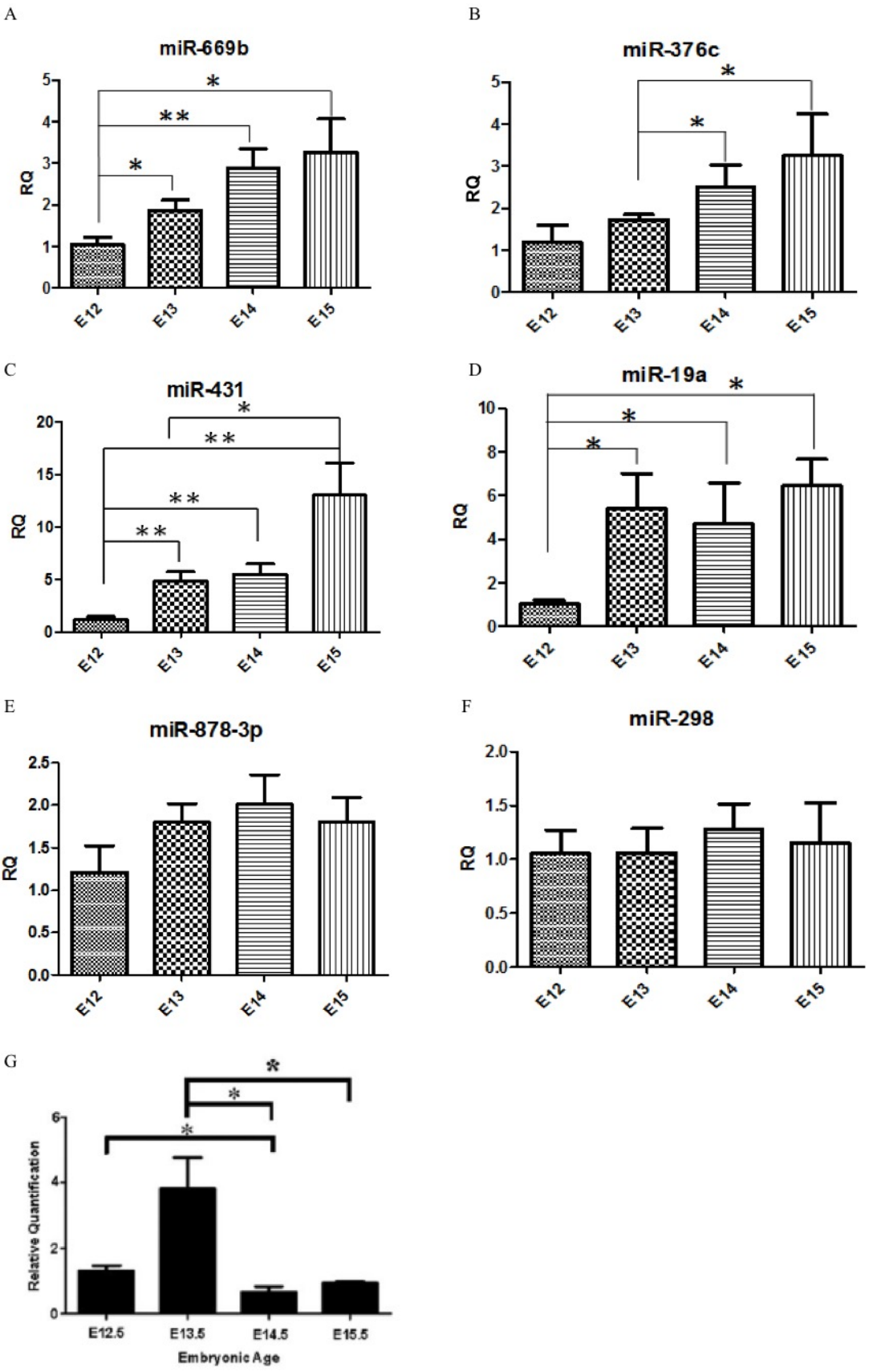

A
miR-669b
RQ
E12
E13
E14
E15
B
miR-376c
C
miR-431
D
miR-19a
E
miR-878-3p
F
miR-298
G
Relative Quantification
E12.5
E13.5
E14.5
E15.5
Embryonic Age

Figure 5.12. miRNA expression in the developing mouse palate. miRNA expression levels in wild-type mouse palate shelves from E12 to E15 were measured by qRT-PCR as described in section 4.2.5, p76. miRNA expressions were relative to snRNA RNU6B. Relative quantitative expression of miRNAs at E12 were normalized to 1. Relative quantitative expression during mouse palate development of miRNA: (A) miR-669b; (B) miR-376c; (C), miR-431; (D) miR-19a; (E) miR-878-3p and (F) miR-298. (G) Expression level of *Hoxa2* gene in developing mouse palate. *Hoxa2* gene expressions were relative to β-actin. Figure taken from Smith *et al*., 2009 with permission. RQ: relative quantity of miRNA expression. Bars represent mean ±SEM, n =3. *$p \leq 0.05$ between bars indicated by brackets. **$p \leq 0.01$ between bars indicated by brackets.

5.2.4 miRNA-669b and miR-376c down regulate *Hoxa2* expression in NIH 3T3 cells.

To investigate whether miR-669b and miR-376c impact *Hoxa2* gene expression, I transfected NIH 3T3 cell line with the respective miRNA mimic. NIH 3T3 cells transfected with miR-669b mimic reduced *Hoxa2* mRNA expression by ~30% after 24h of transfection (Fig. 5.13A). Although miR-376c mimic significantly reduced expression of *Hoxa2* after 24h of transfection compared to mock treated cells (transfection reagent only), it did not reach significance when compared to control miRNA treated cells. These differences were not observed after 48h of transfection with either miR-669b or miR-376c mimic, possibly due to the degradation of the miRNA mimics. Western blot analysis showed both miR-669b and miR-376c mimics down regulated Hoxa2 protein expression after 24h of transfection (Fig. 5.13B).

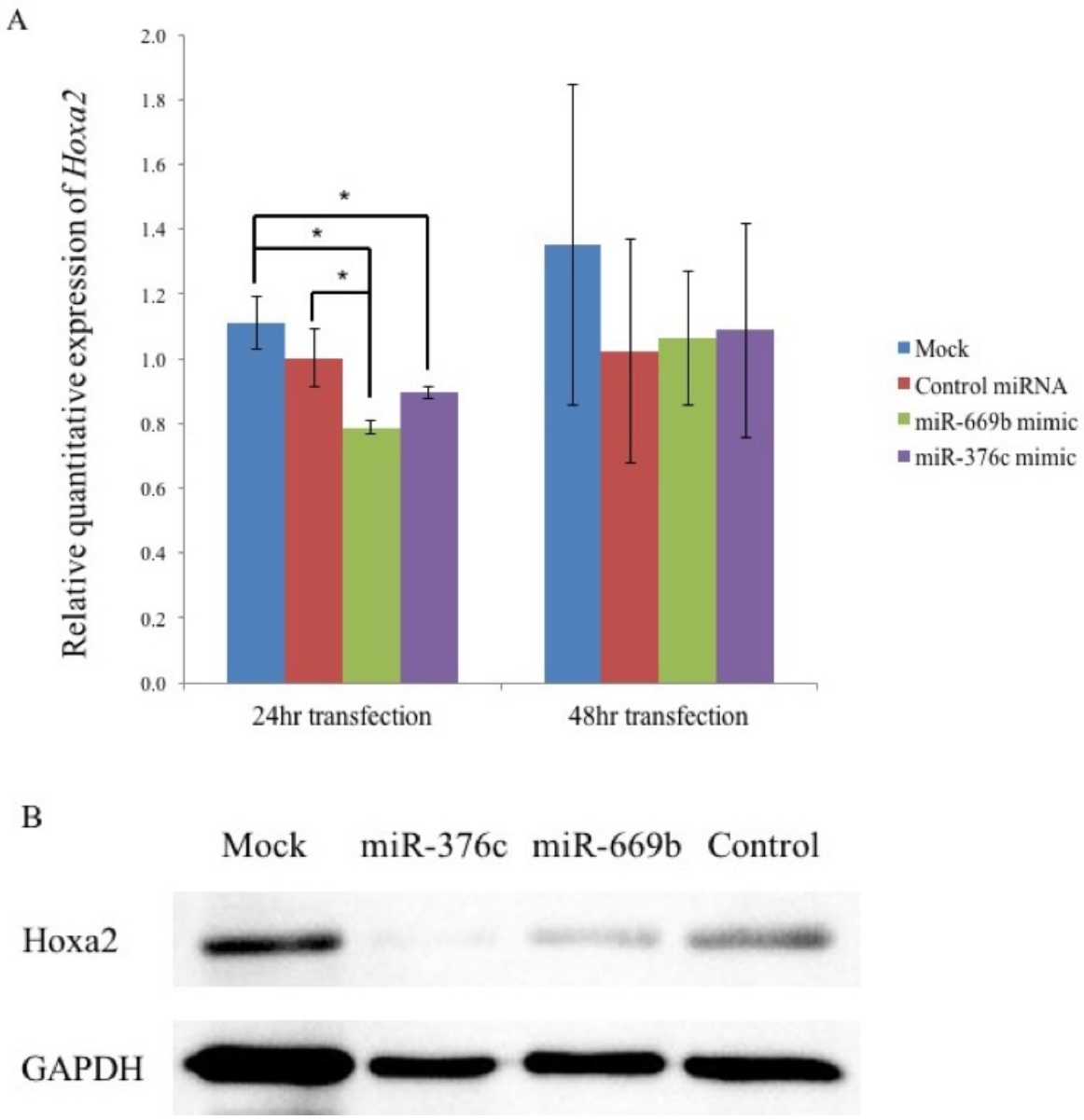


Figure 5.13. Effect of miR-669b and miR-376c mimics on *Hoxa2* expression in NIH 3T3 cells. NIH 3T3 cells were transfected with miR-376c mimic, miR-669b mimic, control miRNA and mock treatment, respectively. (A) After 24h and 48h of transfection, total RNAs were isolated and qRT-PCR was carried out using *Hoxa2* Taqman primers as described in 4.2.5. Bars represent mean ±SEM, n =3. *$p \leq 0.05$ between bars indicated by brackets. (B) NIH 3T3 cells were transfected with miR-376c mimic, miR-669b mimic, control miRNA or mock treatment, and whole cell lysates were collected for western blot analysis after 24h. GAPDH was used as a loading control.

5.2.5 miR-669b binds directly to mouse *Hoxa2* 3'UTR

Since both miR-669b and miR-376c repress *Hoxa2* expression in NIH 3T3 cells (Fig. 5.13), I further investigated whether this effect on *Hoxa2* expression is due to direct binding of miRNAs to *Hoxa2* 3' UTR. To answer this question, I carried out dual-luciferase assays in NIH 3T3 cells. The vectors pEZX-MT01 (Genecopoeia®) encoding *firefly* luciferase (hLuc) and *renilla* luciferase (hRLuc) pEZX-MT01-*Hoxa2* (Genecopoeia®) containing full length mouse *Hoxa2* 3' UTR were used in these experiments. If *Hoxa2* 3' UTR was being targeted by the miRNAs, the translation of *firefly* luciferase will be affected and any change in enzyme activities will be detected (Fig. 4.3). *Renilla* luciferase was used as an internal control to normalize *firefly* luminescence. In the experimental group, pEZX-MT01-*Hoxa2* was co-transfected with miR-669b mimic, miR-376c mimic or control miRNA, respectively. A significantly reduced luminescence signal was observed in cells transfected with miR-669b mimic, indicating miR-669b can bind directly to mouse *Hoxa2* 3' UTR (Fig. 5.14). No significant down regulation of *firefly* luciferase activity was observed in samples transfected with miR-376c (Fig. 5.14). Hence, the down regulation of *Hoxa2* expression in NIH 3T3 cells following transfection of miR-376c mimic does not appear to be through direct binding of miR-376c to *Hoxa2* 3'UTR. For the control group, I used pEZX-MT01 blank vector without *Hoxa2* 3'UTR and found no significant difference in NIH 3T3 cells transfected with miR-669b mimic, miR-376c mimic or control miRNA (Fig. 5.14).

miR-669b has two predicted binding sites on *Hoxa2* 3'UTR and a AA-CG mutation was engineered in the predicted seed sequences (Fig. 5.5A) on pEZX-MT01-*Hoxa2* vector (pEZX-MT01- *Hoxa2* mut1 and pEZX-MT01- *Hoxa2* mut2, Fig. 5.5A). A construct with both predicted binding sites that were mutated was also generated (pEZX-MT01-*Hoxa2* mut1+2). Luciferase

assay showed that mutation at single binding site alone was not sufficent to block the effect of miR-669b on *Hoxa*2 3'UTR, only after both binding sites were mutated, miR-669b mimic could no longer reduce the luciferase activity, suggesting that both predicted binding sites are required for the effective miR-669b function (Fig. 5.15).

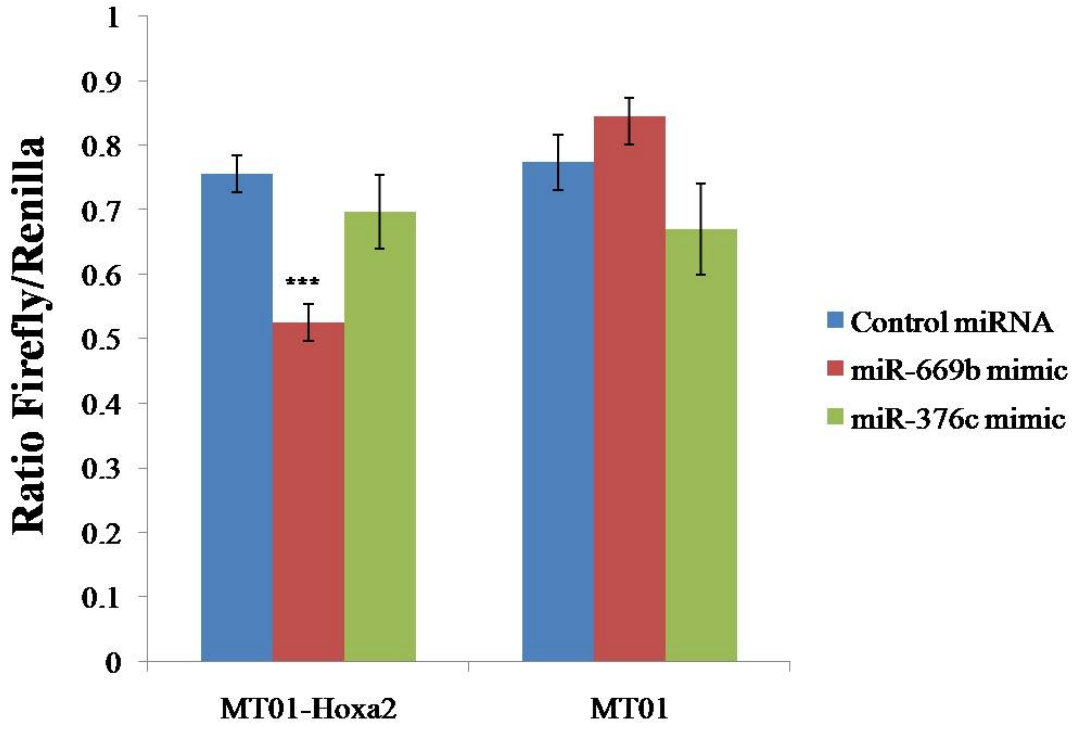


Figure 5.14. Luciferase assay revealed direct interactions between miR-669b and mouse *Hoxa2* 3'UTR. NIH 3T3 cells co-transfected with luciferase vector and miRNA mimics. Luciferase reporter vectors with (pEZX-MT01-*Hoxa2*) or without *Hoxa2* 3'UTR (pEZX-MT01) were co-transfected with miR-669b mimic, miR-376c mimic or control miRNA, respectively *Firefly* luminescence signal was normalized to *renilla* luminescence signal. Y-axis shows ratio of *firefly* and *renilla* luminescence signal. Bars represent mean ±SEM, n =4. ***p ≤ 0.001 compared to control miRNA.

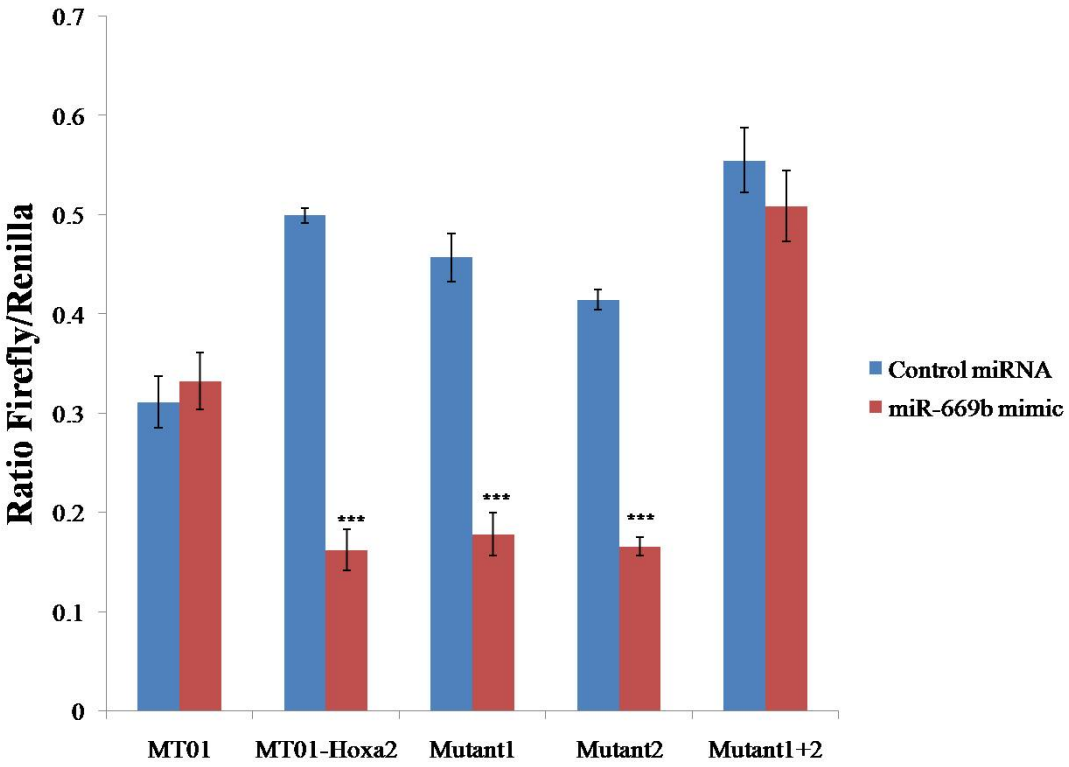


Figure 5.15. Luciferase assay revealed that seed sequence mutations affect miR-669b binding capacity to *Hoxa2* 3' UTR. *Firefly* luminescence signal was normalized to *renilla* luminescence signal. Y-axis shows ratio of *firefly* and *renilla* luminescence signal. Three different vectors with mutations in seed binding sites of *Hoxa2* 3' UTR (pEZX-MT01- *Hoxa2* mut1, pEZX-MT01- *Hoxa2* mut2 and pEZX-MT01-*Hoxa2* mut1+mut2) were tested for the luciferase assay. Bars represent mean ±SEM, n =3. ***p ≤ 0.001 compared to control miRNA .

The findings above show that the overexpression of both miR-669b and miR-376c decreases the expression of *Hoxa2* at transcriptional and translational level. Using luciferase reporter assay, two direct miR-669b binding sites were revealed in mouse *Hoxa2* 3'UTR and both sites proved to be functional. miR-376c does not have direct binding sites on *Hoxa2* 3'UTR. The effect of miR-376c on *Hoxa2* expression may be due to the binding of miR-376c to another part of the *Hoxa2* gene, or due to an indirect effect.

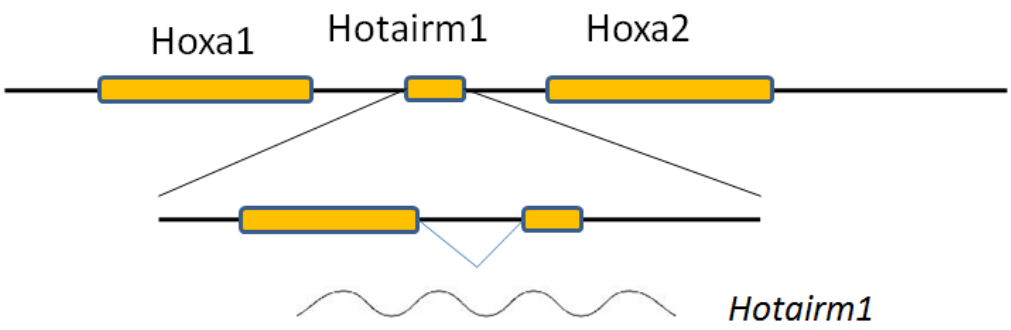
5.3 *mHotairm1* lncRNA regulates *Hoxa1* and *Hoxa2* gene expression via an epigenetic mechanism

HOTAIRM1 is known to activate 3' HOXA genes (HOXA genes located at the 3' end of HOXA loci) in human cells (Zhang *et al*., 2009), yet no similar transcript has been reported in other species and little is known about how 3' HOXA genes are regulated by *HOTAIRM1*. In the following study, I identified the existence of a noncoding transcript in mouse that shares sequence similarity with human *HOTAIRM1*, and referred it as the mouse *Hotairm1* (*mHotairm1*). I further investigated the roles and mechanisms of *mHotairm1* in regulating the expression of two 3' Hoxa genes in mice, namely *Hoxa1* and *Hoxa2*.

5.3.1 Predicted mouse *Hotairm1* sequence analysis

To investigate the existence of *mHotairm1* noncoding transcript in mouse, first I analyzed the human *HOTAIRM1* sequence using BLAST with mouse RefSeq RNA, and found a predicted transcript in mouse that shared some sequence similarity with human *HOTAIRM1* (Fig. 5.16C). This transcript in mouse is transcribed between *Hoxa1* and *Hoxa2* (Fig. 5.16A), and it is 522 nt long (Fig. 5.16B) with two exons and one intron that undergoes splicing during transcription (Fig. 5.16D). Next, three pairs of overlapping primers were designed (Fig. 4.4, Table 4.3) and PCR was carried out for a cDNA library that was generated from NIH 3T3 cells. The amplied DNA fragments were sequenced and a 462 nt sequence was identified, hereafter I refered it as *mHotairm1* [Fig. 5.16B (sequence in orange), Appendix 7-9].

A



B

CGGCCGCTCCCGGAGCTGACTTGGAGCACTGGGACCAAAGGGAGTCGAGACTGC
CTTCTGCGCGCGCCCGGCTTTGCGCGCCTCCGCCACCAGATGTGGGGGGATGGGA
GGCCCCCTCCGCGGCCCCTTCCCCACCCAGCCCAGAAAGCTGAACTGGCAAGAG
GTCTGTTTTTCCTGAACCCATCCACAGCTGGGAGATTAATCAACCACACTGAAAAT
GGGGGTGTGGGGGAGGGAAAGGAAGAGTTGGAACGTAGATGTTTGAAACAAATG
TGTATAAATAAATGAATTTTTGATAACTCCGTTATTGACCTAGAAACTAGCAGCTTG
GTAAGGGAACTCCATTCCACTCCACTCGTCCTAGAACTGGAAGTTTTTGTAGGCA
CTTTTCCTCTCCACACTCAAAAGCTTGGGCTAGGGCCAACTCAGGCTGCCCAAGC
CCATTTCTATTACTAATGTAACTCTATGGCCTGAGTCTCAACACTGAAAACCAAATT
CATTCCCTTAGGGGGGAAAAATCCA

C. Sequence similarity between human and mouse *Hotairm1*.

```
Mouse   51   CTGCCTTCTGCGCGCGCCCGGCTT---TGCGCGCCTCCGCCA-----CCAGATGTGGGGG  102
             ||||  ||||||||||||||| ||    ||| |||| ||||||     || |  |  ||||
Human   173  CTGCGTTCTGCGCGCGCCCGACTCCGCTGCCCGCC-CCGCCAGGCCTCCGGGAGGTGGGG  231

Mouse   103  GATGGGAGGCCCCCTCCGCG---GCCCCTTCCCCACCCAGCCCAGAAAGCTGAACTGGCA  159
             | ||||||||  || ||||    ||||| ||||||||   |   ||||| |||||||||
Human   232  GCTGGGAGGCGTCCCCCGCTCCCGCCCCCTCCCCACCGTTCAATGAAAGATGAACTGGCG  291

Mouse   160  AGAGGTCTGTTTTTCCTGAACCCATCCACAGCTGGGAGATTAATCAACCACACTGAAAAT  219
             ||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||||
Human   292  AGAGGTCTGTTTTGCCTGAACCCATCAACAGCTGGGAGATTAATCAACCACACTGAAAAT  351

Mouse   220  GGGG------GTGTGGGGGAGGGAAAGGAAGAGTTGGAACGTAGATGTTTGAAACAAATG  273
             | ||       | |||||||||        | |||| || || | |||||||||||| |
Human   352  GTGGAGGGATTTATGGGGGAGG--------GGGTTGAAATGTGGGTGTTTGAAACAAAAG  403
```

```
Mouse  274  TGTATAAATAAATGAATTTTTGATAACTCCGTTATTGACCTAGAAACTAGCAGCTTGGTA  333
            |||||||| ||||||||| |||||||||  |||||||||| || ||| | |||||  ||
Human  404  TGTATAAACAAATGAATTGTTGATAACTTAGTTATTGACCTGGAGACTGGTAGCTTATTA  463

Mouse  334  AGGGAACTCCATTCCACTCCACTCGTCCTAGAACTGGAAGTTTTTGTAGGCACTTTTCCT  393
            | | |||||| |   ||||      |||| ||  |||  |||| ||||||||||||   |
Human  464  AAGAAACTCCGTGTTACTC----ATTCCTGGAGTTGGGGGTTTCTGTAGGCACTTTATTT  519

Mouse  394  CTCCACACTCAAAAGCTTGGGCTAGGGCCAA--CTCAGGCTGCCCAAGCCCATTTCTATT  451
            ||||||  |||| |||||||||| || ||||  || || ||| ||||  |    ||||||
Human  520  CTCCACTTTCAAGAGCTTGGGCTTGGCCCAAATCTTAGACTGTCCAATTCTGCCTCTATT  579

Mouse  452  ACTAATGTAACTCTATGGCCTGAGTCTCAACACTGAAAACCAAAT  496
            || ||| ||| |||||||| |||  ||   ||||||||| |||||
Human  580  ACCAATTTAAATCTATGGCTTGAACCTGTGCACTGAAAATCAAAT  624
```

D. Comparation of *mHotairm1* RNA sequence to mouse genome.

(a)

```
mHotairm1  162         AGGTCTGTTTTTCCTGAACCCATCCACAGCTGGGAGATTAATCAACCACACTGAAAATGG 221
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52161929  AGGTCTGTTTTTCCTGAACCCATCCACAGCTGGGAGATTAATCAACCACACTGAAAATGG 52161988

mHotairm1  222         GGGTGTGGGGGAGGGAAAGGAAGAGTTGGAACGTAGATGTTTGAAACAAATGTGTATAAA  281
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52161989  GGGTGTGGGGGAGGGAAAGGAAGAGTTGGAACGTAGATGTTTGAAACAAATGTGTATAAA 52162048

mHotairm1  282         TAAATGAATTTTTGATAACTCCGTTATTGACCTAGAAACTAGCAGCTTGGTAAGGGAACT  341
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52162049  TAAATGAATTTTTGATAACTCCGTTATTGACCTAGAAACTAGCAGCTTGGTAAGGGAACT 52162108

mHotairm1  342         CCATTCCACTCCACTCGTCCTAGAACTGGAAGTTTTTGTAGGCACTTTTCCTCTCCACAC  401
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52162109  CCATTCCACTCCACTCGTCCTAGAACTGGAAGTTTTTGTAGGCACTTTTCCTCTCCACAC 52162168

mHotairm1  402         TCAAAAGCTTGGGCTAGGGCCAACTCAGGCTGCCCAAGCCCATTTCTATTACTAATGTAA  461
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52162169  TCAAAAGCTTGGGCTAGGGCCAACTCAGGCTGCCCAAGCCCATTTCTATTACTAATGTAA 52162228

mHotairm1  462         CTCTATGGCCTGAGTCTCAACACTGAAAACCAAATTCATTCCCTTAGGGGGGAAAAATCC  521
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52162229  CTCTATGGCCTGAGTCTCAACACTGAAAACCAAATTCATTCCCTTAGGGGGGAAAAATCC 52162288
```

```
mHotairm1  522          A  522
                        |
Mouse genome  52162289  A  52162289
```

（b）

```
mHotairm1  1           CGGCCGCTCCCGGAGCTGACTTGGAGCACTGGGACCAAAGGGAGTCGAGACTGCCTTCTG  60
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52158524  CGGCCGCTCCCGGAGCTGACTTGGAGCACTGGGACCAAAGGGAGTCGAGACTGCCTTCTG 52158583

mHotairm1  61          CGCGCGCCCGGCTTTGCGCGCCTCCGCCACCAGATGTGGGGGGATGGGAGGCCCCCTCCG  120
                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52158584  CGCGCGCCCGGCTTTGCGCGCCTCCGCCACCAGATGTGGGGGGATGGGAGGCCCCCTCCG 52158643

mHotairm1  121         CGGCCCCTTCCCCACCCAGCCCAGAAAGCTGAACTGGCAAGAGGT  165
                       |||||||||||||||||||||||||||||||||||||||||||||
Mouse genome 52158644  CGGCCCCTTCCCCACCCAGCCCAGAAAGCTGAACTGGCAAGAGGT 52158688
```

Figure 5.16. Sequence analysis of mouse *Hotairm1.* A. Mouse *Hotairm1* is transcribed between *Hoxa1* and *Hoxa2*. It has two exons, and one intron that is spliced during transcription. B. The 522 nt predicted transcript in mouse that shared sequence similarity with human *HOTAIRM1* was listed. Sequence identified by cloning and sequencing of mouse *Hotairm1* transcript (462 nt) is shown in orange. Splicing happen at nucleotide 163, with AG (marked in blue) the 5' splicing signal and G (marked in green) the 3' splicing signal. C. Analysis of sequence similarity between human and mouse *Hotairm1*. *mHotairm1* RNA sequence was blasted with human *HOTAIRM1* RNA sequence. D. mHotairm1 RNA sequence was blasted with mouse genomic DNA. Two ranges were found (a) and (b), indicating that *mHotairm1* is spliced.

5.3.2 Expression of *mHotairm1* in mouse tissues and cell lines

To further confirm the existence of *mHotairm1* and to study its function, I investigated the expression of *mHotairm1* in mouse embryonic tissues. Also, I investigated whether the expression of *mHotairm1* in NIH 3T3 cells is induced by an activator of *Hox* gene expression, namely, the all-*trans* retinoic acid (ATRA) (Bertani *et al*., 2011).

5.3.2.1 *mHotairm1* is expressed in several different mouse tissues

The expression of *mHotairm1* was examined in the mouse developing palate using real-time PCR. In the mouse, palatogenesis occurs between E12 to E15 (Smith *et al*., 2009; 2013) and my findings showed that *mHotairm1* expression exhibited the highest expression at E12, and gradually decreased with the lowest expression at E15 (Fig. 5.17A). *mHotairm1* is also expressed in the head, forelimbs, hindlimbs and tail region of E13 embryos with the highest expression in the tail (Fig. 5.17B).

5.3.2.2 *In situ* hybridization histochemistry of *mHotairm1* in mouse palate

To determine precisely where the *mHotairm1* transcript is being expressed in the developing mouse palate, I utilized *in situ* hybridization histochemistry (ISH). The *mHotairm1* transcript appeared to be primarily expressed in the medial edge epithelial cells of the developing palate at E14 and E15 (Fig. 5.18). The location of its expression at the seam of the medial edge epithelia prior to palatal fusion may indicate its role in facilitating palatal fusion. Further experiments will need to be carried out to identify the role of *mHotairm1* in palatal development.

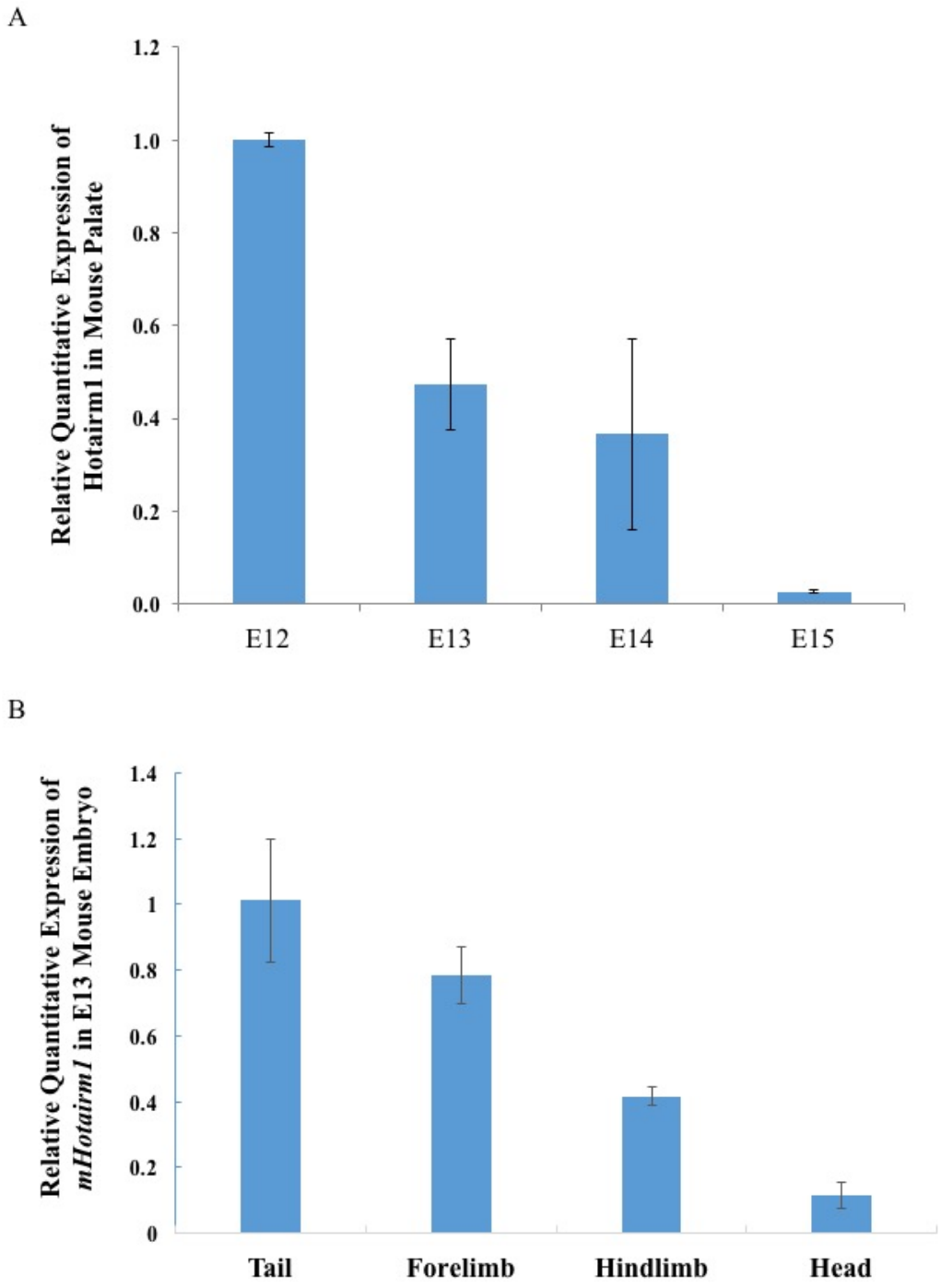


Figure 5.17. Mouse *Hotairm1* expression in the developing palate (E12 to E15) and mouse tissues from E13 embryos. A. *mHotairm1* expression levels in wild-type mouse palatal shelves from E12 to E15 were quantitated with qRT-PCR. B. *mHotairm1* expression detected in head, forelimbs, hindlimbs and tail region from E13 embryos by qRT-PCR. Bars represent mean ±SEM, n =3.

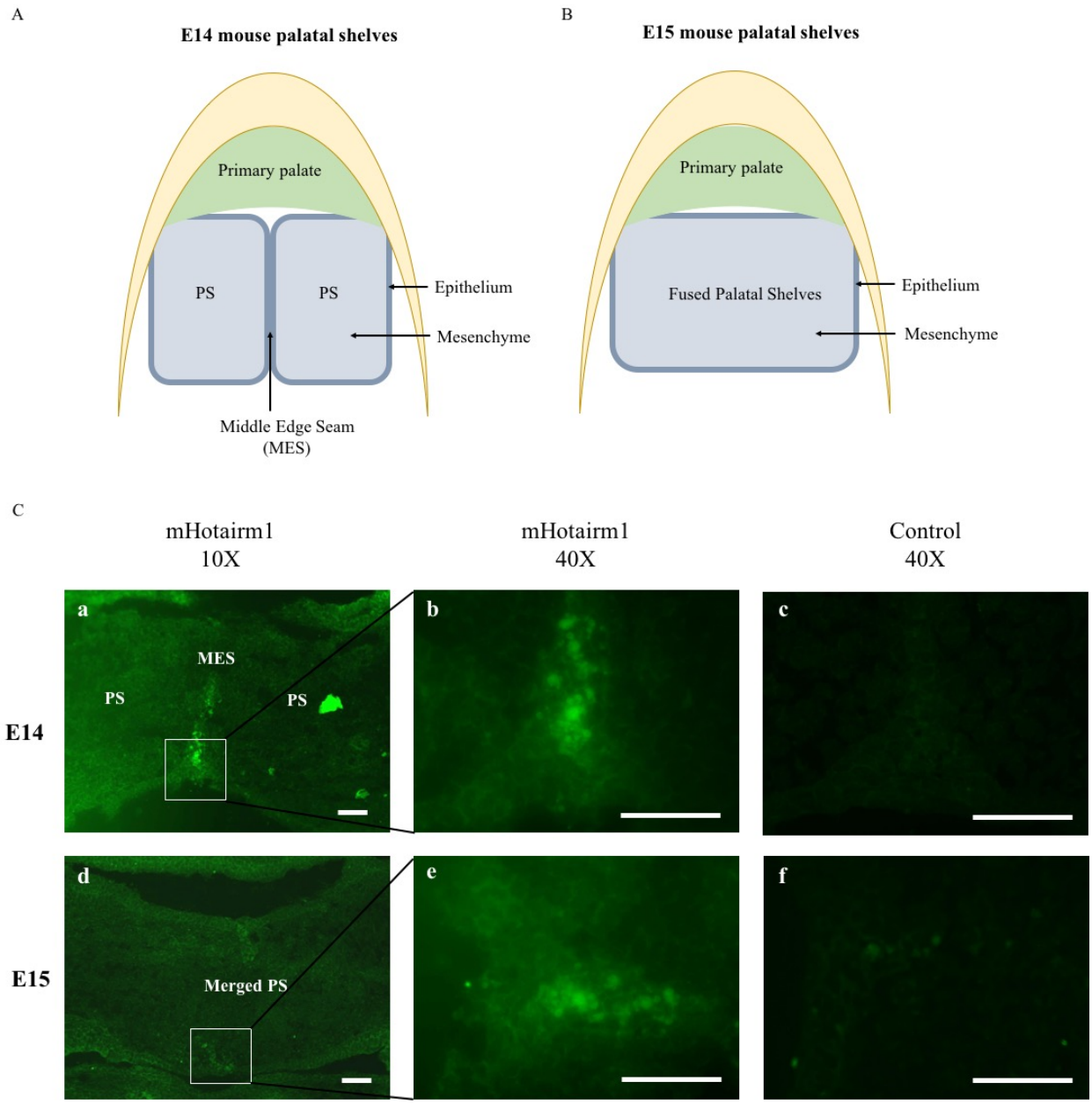


Figure 5.18. *In situ* hybridization histochemistry of *mHotairm1* in mouse palate. A-B: Schematic diagram of E14 and E15 mouse palatal shelvies. PS: palatal shelvies. MES: middle edge seam. C: Distribution of *mHotairm1* in E14 and E15 mouse palate detected by *in situ* hybridization histochemistry. Green staining represents the expression of *mHotairm1* transcript at E 14 (a,b), and E15 (d,e). (c,f): control in situ probes were used and no staining was observed. Scale bar indicate 5 microns.

5.3.2.3 ATRA induces *mHotairm1* expression

ATRA is a known inducer for the expression of *Hox* genes and has also been reported to induce lncRNA transcription (Bertani *et al*., 2011). To determine how *mHotairm1* is regulated, NIH 3T3 cells were treated with ATRA. Results showed that the expression of *mHotairm1* as well as that of *Hoxa1* and *Hoxa2* were significantly increased after 24 h treatment with ATRA (Fig. 5.19). These results confirmed that *mHotairm1* is a mouse noncoding transcript that is present in mouse embryonic tissues and that its expression can be induced by ATRA in the NIH 3T3 cell line

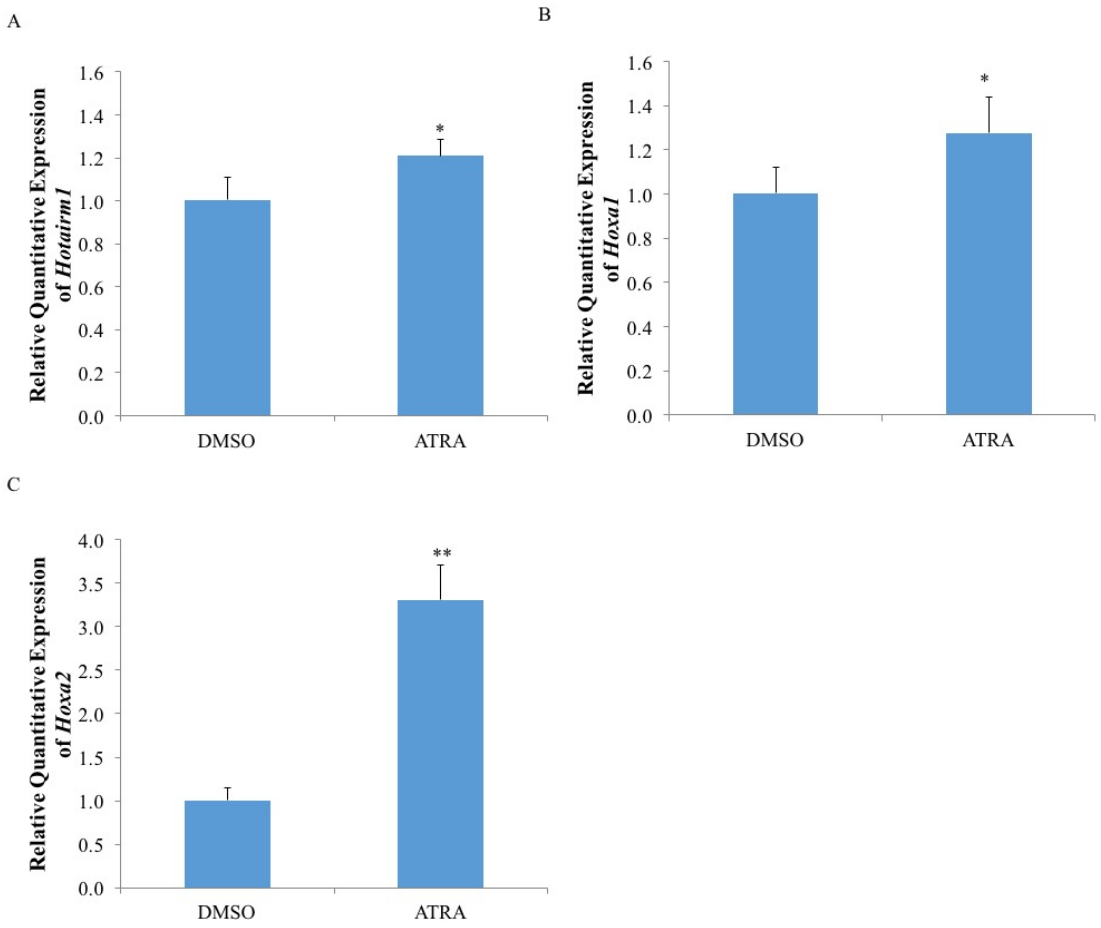


Figure 5.19. All-*trans* retinoic acid induces *mHotairm1*, *Hoxa1* and *Hoxa2* expression. NIH 3T3 cells were treated with 10^{-6}M ATRA. After 24h, RNA was isolated and *mHotairm1* (A), *Hoxa1* (B) and *Hoxa2* (C) expression were quantified using qRT-PCR. Bars represent mean ±SEM, n =3. *p ≤ 0.05, **p ≤ 0.01 compared to DMSO treated control group.

These findings show that *mHotairm1* is expressed differently in different mouse embryonic tissues. It is also expressed in NIH 3T3 cells and its expression can be induced by ATRA, which agrees with the findings of Zhang *et al* (2009) that human HOTAIRM1 can be induced by ATRA in myeloid differentiation.

5.3.3 *mHotairm1* regulates *Hoxa1* and *Hoxa2* expression via histone methylation

The confirmation of the existence of *mHotairm1* led to my next research question: what is the function of this lncRNA? The human *HOTAIRM1* is known to regulate 3′ HOXA genes, and several lncRNAs transcribed within Hox cluster genes are also known to regulate expression of nearby *Hox* genes (Wang *et al*., 2011, Bertani *et al*., 2011). Hence, I further investigated whether *mHotairm1* regulates the expression of nearby *Hoxa1* and *Hoxa2* genes and what if any, were regulatory mechanisms involved.

5.3.3.1 Knockdown expression of *mHotairm1* leads to decreased expression of *Hoxa1* and *Hoxa2*

To determine what impact *mHotairm1* had on mouse *Hoxa1* and *Hoxa2* expression, NIH 3T3 cells were transfected with *mHotairm1* siRNA. Exposure to *mHotairm1* siRNA decreased *mHotairm1* expression in NIH 3T3 cells after 72h, and resulted in significantly decreased expression of both *Hoxa1* and *Hoxa2* genes (Fig. 5.20).

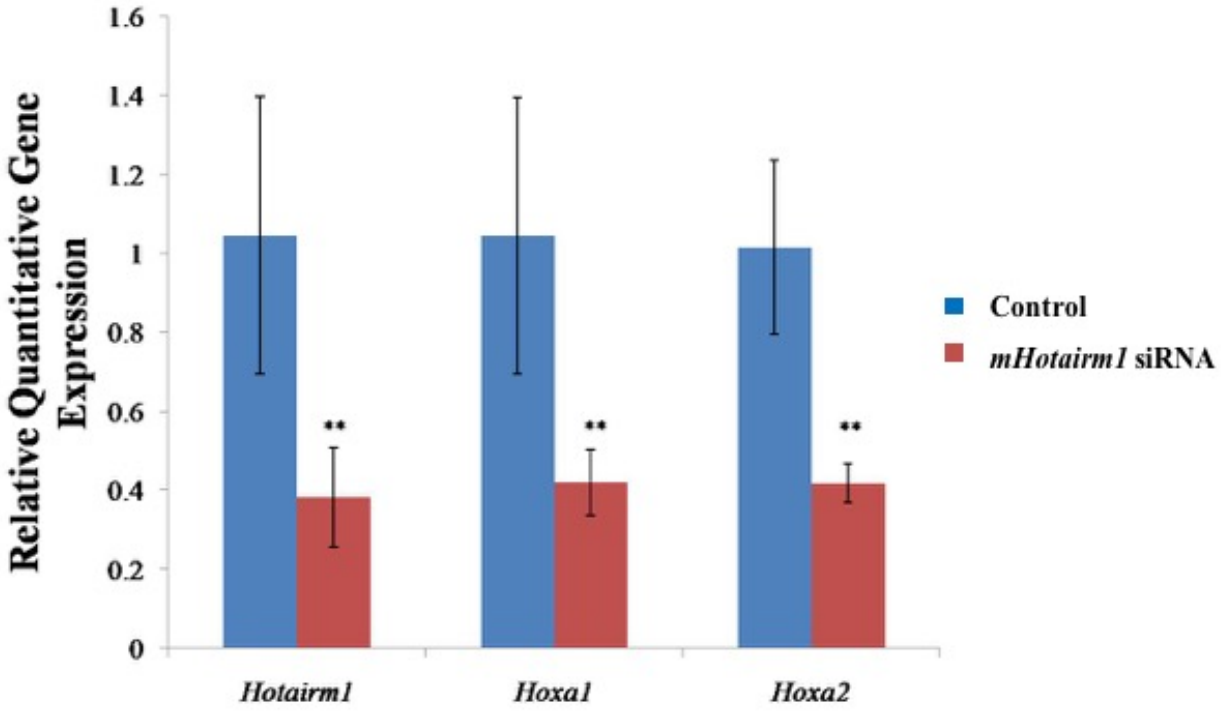


Figure 5.20. Knockdown of *mHotairm1* leads to decreased expression of *Hoxa1* and *Hoxa2*. NIH 3T3 cells were transfected with 150 ng of *mHotairm1* siRNA and control siRNA, respectively. After 72 h of transfection, total RNAs were isolated, reverse transcribed and quantified using qRT-PCR with *Hoxa1*, *Hoxa2* and *mHotairm1* primers. Bars represent mean ±SEM, n =3. **p ≤ 0.01 compared to respective control.

5.3.3.2 *mHotairm1* interacts with promoters of 3' HoxA genes and WDR5.

To determine the potential mechanism of how blocking the expression of *mHotairm1* causes the down expression of *Hoxa1* and *Hoxa2* expression, I carried out a "CHART" experiment (Simon *et al*., 2011) to investigate any interactions between *mHotairm1*, *Hoxa1* and *Hoxa2* promoter and the H3K4 methyltransferase complex MLL1/WDR5. DNA samples bound to *mHotaim1* probes were collected from the "CHART" experiment and tested using PCR. Probe 1, 2, 3 were designed to detect *mHotairm1* and Probe C (anti-sense to Probe 2) was used as the control probe (Fig. 4.4, Table 4.3). Probe 2 showed much stronger interaction with *mHotairm1* (Fig. 5.21 and Fig. 5.22) compared to the control probe, indicating a specific recognition of

mHotairm1 by Probe 2. PCR results from the DNA fragments pulled down by Probe 2 showed a strong interaction between *mHotairm1* and the promoters of 3' HoxA genes (*Hoxa1, Hoxa2, Hoxa3 and Hoxa5*) but not with 5' HoxA gene (*Hoxa13*) when compared to Probe C (Fig. 5.21). Protein samples pulled down by Probe 2 also showed strong interaction between *mHotairm1* and WDR5, a subunit in TrxG complex (Fig. 5.22A). No interaction was observed between *mHotairm1* and CBP (CREB-binding protein), a histone acetyltransferase. These results indicate that there is strong interaction between *mHotairm1* and promoters of 3' HoxA genes including *Hoxa1* and *Hoxa2*, and the TrxG complex MLL1/WDR5.

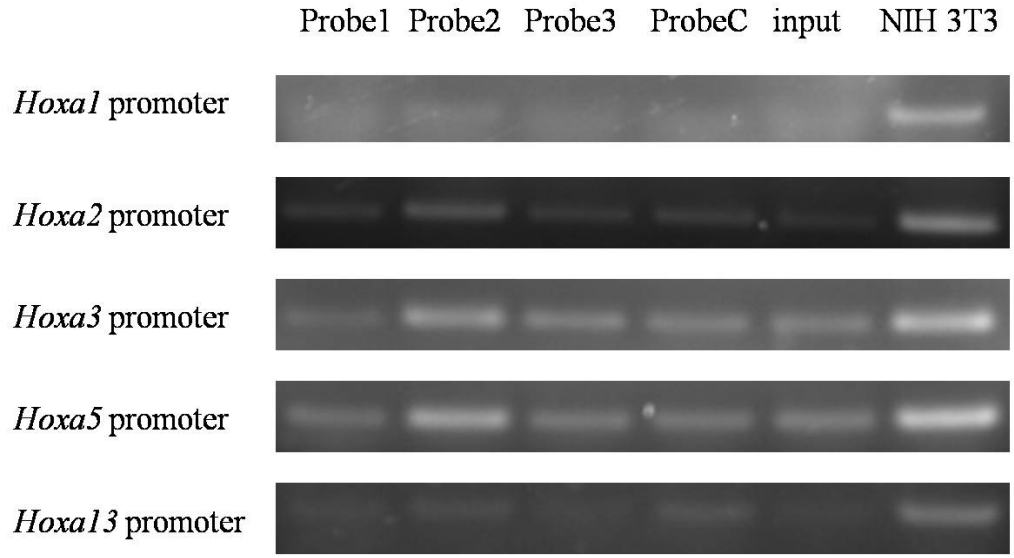


Figure 5.21. PCR amplification of CHART enriched DNA fragments. DNA fragments that can interact with *mHotairm1* were pulled down together with *mHotairm1* using gene specific Probe 1, 2, 3. Probe C was used as control. Each pull down sample was PCR tested with *Hoxa1, Hoxa2, Hoxa3, Hoxa5* and *Hoxa13* promoter primers (see Table 4.3 for probe and primer sequences).

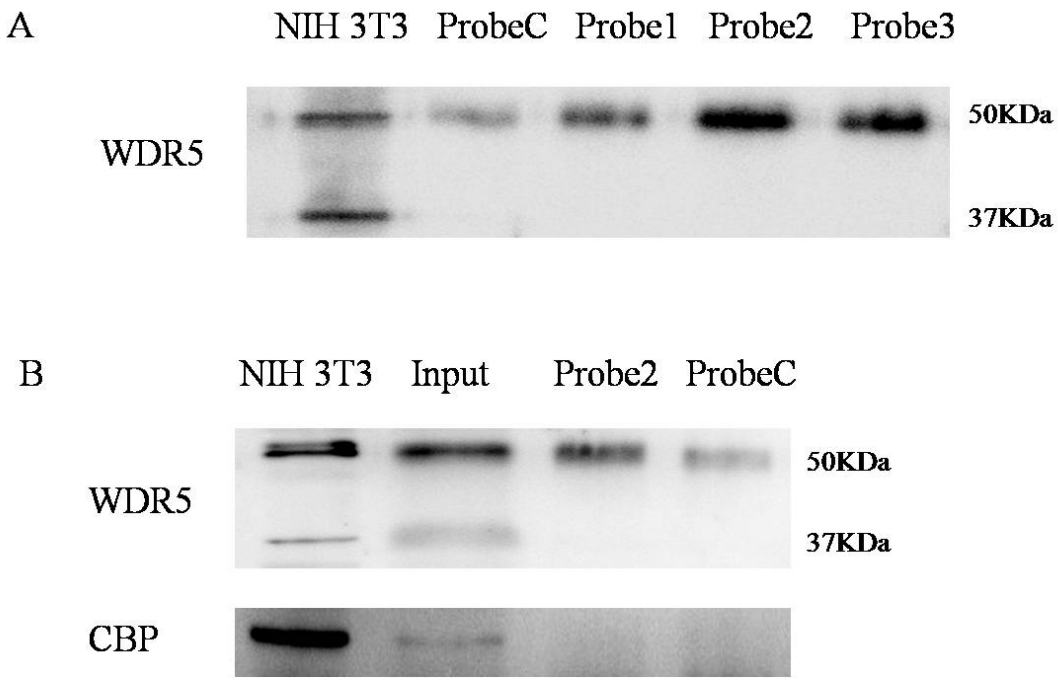


Figure 5.22. Western blot of CHART enriched protein samples. A. Protein samples interacting with *mHotairm1* was pulled down together with it using gene specific Probe 1, 2, 3. Probe C was used as control. Each pull down sample was tested with WDR5 antibody in western blot. NIH 3T3 cell lysate was used as a positive control. Two WDR5 bands were observed in NIH 3T3 cell lysate, a WDR5 band at ~37KDa and a modified WDR5 band at ~ 50KDa. Only ~50KDa bands were seen in CHART enriched protein samples. B. No CBP band was detected in western blot in samples pulled down by *mHotairm1* specific probe 2.

5.3.3.3 *mHotairm1* binds directly to MLL1/WDR5 complex.

5.3.3.3.1 Construction of GST fusion protein vectors.

To further confirm the interaction between *mHotairm1* and MLL1/WDR5 complex, GST pull down assay was used. WDR5 and MLL1 (amino acid 3810-3963) coding cDNAs were successfully amplified from NIH 3T3 cDNA samples, and cloned into T-easy vectors and further confirmed with sequencing (Fig. 5.23, Appendix 10-11). From T-easy vectors, the two fragments were then cloned into pGEX-6p-1 vector that contains a GST tag. MLL1 (3810-3963), WDR5 and GST protein expression were successfully induced with IPTG in DH5α *E. coli* (Fig. 5.24A-C). All GST fused proteins were purified with Glutathione agarose beads. Single bands were observed in purified samples (Fig. 5.24D).

5.3.3.3.2 Binding of *mHotairm1* to MLL1/WDR5 complex

To confirm the interaction between mHotairm1 and MLL1/WDR5 complex, I carried out GST pull down experiments. Using GST fused MLL1 (3810-3963) and WDR5 proteins, the RNA pulled down were converted into cDNA and amplified with *mHotairm1* primers mush138 F and mush138 R (Table 4.3). PCR results indicated that *mHotairm1*can be pulled down with both GST-MLL1 (3810-3963) and GST-WDR5, but not with control GST (Fig. 5.24E).

A. DNA and protein sequence of full length mouse WDR5

```
ATG GCC ACA GAG GAG AAG AAG CCA GAG ACA GAG GCT GCA AGA GCA CAG
 M   A   T   E   E   K   K   P   E   T   E   A   A   R   A   Q
CCC ACT CCT TCC TCA TCA GCC ACA CAG AGC AAG CCC ACA CCA GTT AAG
 P   T   P   S   S   S   A   T   Q   S   K   P   T   P   V   K
CCA AAC TAT GCC CTG AAG TTC ACC CTG GCT GGC CAC ACC AAA GCT GTG
 P   N   Y   A   L   K   F   T   L   A   G   H   T   K   A   V
TCC TCT GTG AAG TTC AGC CCC AAT GGG GAA TGG TTG GCA AGT TCA TCT
 S   S   V   K   F   S   P   N   G   E   W   L   A   S   S   S
GCT GAT AAA CTC ATT AAA ATT TGG GGA GCA TAT GAT GGA AAG TTT GAG
 A   D   K   L   I   K   I   W   G   A   Y   D   G   K   F   E
AAA ACT ATA TCT GGT CAC AAA CTG GGA ATA TCT GAT GTA GCG TGG TCA
 K   T   I   S   G   H   K   L   G   I   S   D   V   A   W   S
TCA GAT TCT AAC CTC CTT GTG TCT GCC TCT GAT GAT AAA ACT TTG AAG
 S   D   S   N   L   L   V   S   A   S   D   D   K   T   L   K
ATT TGG GAC GTG AGT TCC GGG AAG TGT CTG AAG ACC CTG AAG GGC CAC
 I   W   D   V   S   S   G   K   C   L   K   T   L   K   G   H
AGT AAC TAC GTC TTC TGC TGC AAC TTC AAC CCC CAG TCC AAC CTC ATC
 S   N   Y   V   F   C   C   N   F   N   P   Q   S   N   L   I
GTC TCA GGG TCT TTT GAT GAA AGT GTG AGG ATA TGG GAC GTG AAG ACA
 V   S   G   S   F   D   E   S   V   R   I   W   D   V   K   T
GGG AAG TGC CTC AAG ACT TTG CCT GCC CAT TCG GAC CCA GTC TCA GCC
 G   K   C   L   K   T   L   P   A   H   S   D   P   V   S   A
GTT CAT TTC AAC CGT GAT GGA TCA TTG ATT GTT TCC AGT AGC TAT GAT
 V   H   F   N   R   D   G   S   L   I   V   S   S   S   Y   D
GGC CTC TGC CGA ATC TGG GAC ACC GCC TCT GGC CAG TGT CTG AAG ACA
 G   L   C   R   I   W   D   T   A   S   G   Q   C   L   K   T
CTC ATT GAT GAT GAC AAT CCT CCA GTG TCC TTC GTG AAG TTC TCT CCA
 L   I   D   D   D   N   P   P   V   S   F   V   K   F   S   P
AAT GGC AAA TAC ATC CTG GCT GCA ACT TTG GAC AAC ACA CTG AAG CTC
 N   G   K   Y   I   L   A   A   T   L   D   N   T   L   K   L
TGG GAC TAC AGC AAG GGG AAG TGC CTG AAG ACA TAC ACT GGC CAC AAG
 W   D   Y   S   K   G   K   C   L   K   T   Y   T   G   H   K
AAT GAG AAG TAC TGC ATA TTT GCC AAC TTC TCC GTG ACA GGC GGG AAG
 N   E   K   Y   C   I   F   A   N   F   S   V   T   G   G   K
TGG ATT GTG TCT GGT TCT GAA GAT AAC CTG GTG TAT ATC TGG AAT CTG
 W   I   V   S   G   S   E   D   N   L   V   Y   I   W   N   L
CAG ACC AAG GAG ATT GTG CAG AAG TTG CAG GGT CAC ACA GAT GTT GTG
 Q   T   K   E   I   V   Q   K   L   Q   G   H   T   D   V   V
ATT TCC ACG GCT TGT CAC CCG ACA GAG AAC ATC ATT GCC TCA GCA GCG
 I   S   T   A   C   H   P   T   E   N   I   I   A   S   A   A
TTA GAG AAC GAC AAA ACA ATC AAA CTG TGG AAG AGT GAC TGC TAA
 L   E   N   D   K   T   I   K   L   W   K   S   D   C
```

B. DNA and protein sequence of mouse MLL1 (amino acid 3810-3963)

```
ATG CCC ATG AGA TTC CGG CAC TTG AAG AAG ACT TCT AAG GAG GCG GTT
 M   P   M   R   F   R   H   L   K   K   T   S   K   E   A   V
GGT GTC TAC AGG TCT CCC ATC CAT GGT CGG GGT CTT TTC TGT AAG AGA
 G   V   Y   R   S   P   I   H   G   R   G   L   F   C   K   R
AAC ATC GAT GCA GGA GAG ATG GTG ATT GAA TAC GCC GGC AAC GTC ATC
 N   I   D   A   G   E   M   V   I   E   Y   A   G   N   V   I
CGC TCC ATC CAG ACA GAC AAG CGT GAG AAG TAC TAT GAC AGC AAG GGC
 R   S   I   Q   T   D   K   R   E   K   Y   Y   D   S   K   G
ATT GGT TGC TAC ATG TTC CGA ATT GAT GAC TCG GAG GTA GTG GAT GCC
 I   G   C   Y   M   F   R   I   D   D   S   E   V   V   D   A
ACC ATG CAT GGA AAT GCT GCA CGC TTC ATC AAT CAC TCT TGT GAG CCT
 T   M   H   G   N   A   A   R   F   I   N   H   S   C   E   P
AAC TGC TAC TCC CGG GTC ATC AAT ATT GAT GGG CAG AAG CAC ATT GTC
 N   C   Y   S   R   V   I   N   I   D   G   Q   K   H   I   V
ATC TTC GCC ATG CGT AAG ATC TAC CGG GGG GAG GAG CTC ACC TAT GAC
 I   F   A   M   R   K   I   Y   R   G   E   E   L   T   Y   D
TAT AAG TTC CCC ATT GAG GAC GCC AGC AAC AAG CTA CCC TGC AAC TGT
 Y   K   F   P   I   E   D   A   S   N   K   L   P   C   N   C
GGC GCC AAA AAA TGC CGC AAG TTC CTG AAC TAA
 G   A   K   K   C   R   K   F   L   N
```

Figure 5.23. Coding sequence of WDR5 and MLL1 (amino acid 3810-3963). DNA seuqences were grouped in triplets with correlating ammino acids marked under DNA sequences in bold letters. DNA sequencing confirmed successful cloing of WDR5 (A) and MLL1 (amino acid 3810-3963) (B).

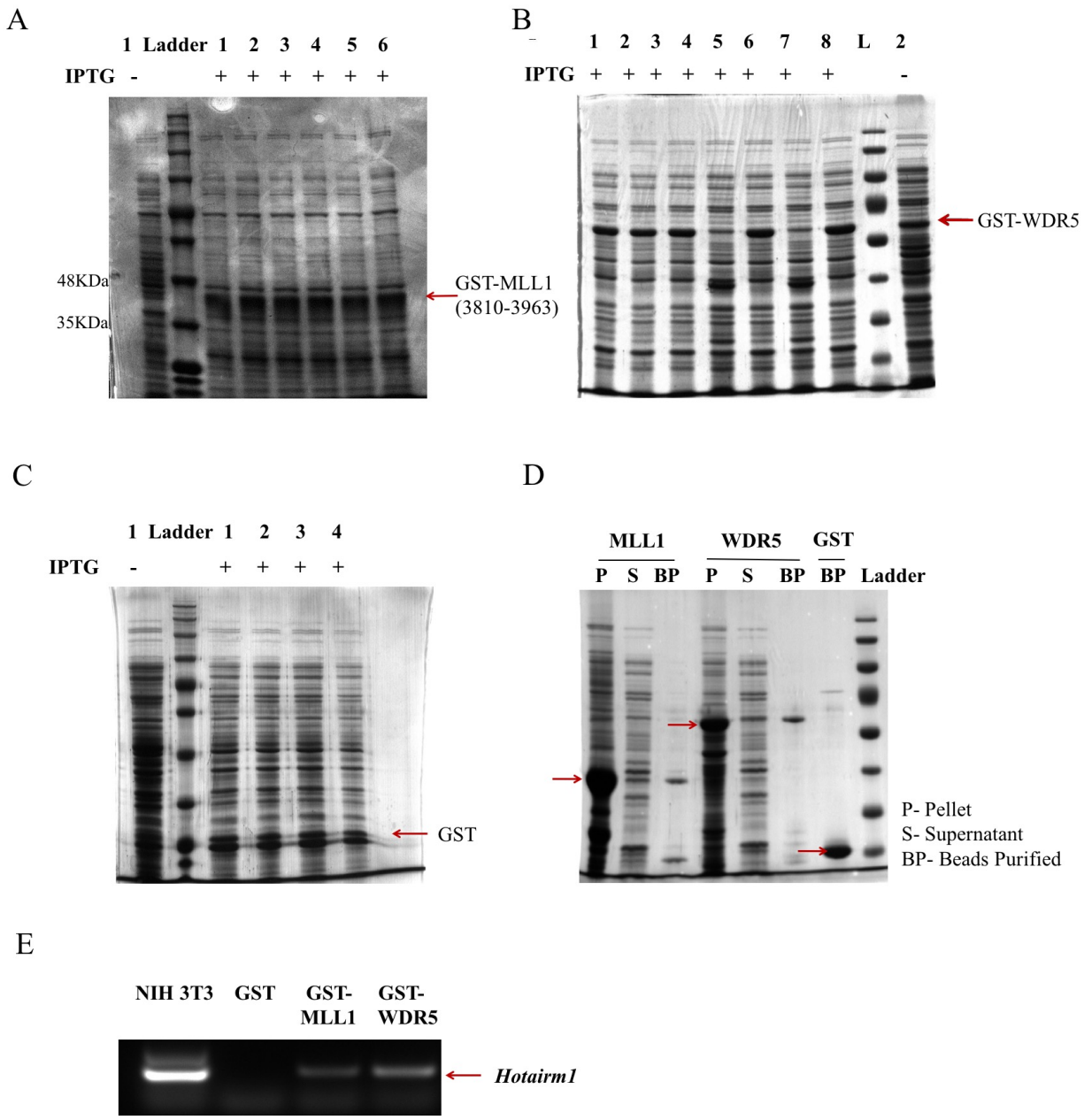


Figure 5.24. GST fusion protein pull down of *mHotairm1*. (A-C) Coomassie blue staining show expressions of GST and GST fused MLL1 (3810-3963) and WDR5 induced with IPTG in DH5α (Pointed with red arrow). Lane numbers indicate different clones. (D) GST fusion proteins were purified with Glutathione agarose beads. P- protein samples from the pellet of *E coli* cell lysate, S- protein samples from the supernatant of *E coli* cell lysate, BP- beads purified protein sample from the supernatant of *E coli* cell lysate. Target bands are pointed with red arrow. (E) PCR amplicant of *mHotairm1* (pointed with red arrow) from GST fusion protein pull down. Bands in MLL1 (3810-3963) and WDR5 pull down indicate direct interaction between *mHotairm1* and MLL1/WDR5 complex. NIH 3T3 cDNA sample was used as a positive control.

5.3.3.4 *mHotairm1* can affect H3K4me3 and H3K27me3 marks on *Hoxa1* and *Hoxa2* chromatin.

Results above confirmed an interaction between *mHotairm1* and TrxG MLL1/WDR5 complex, next, I wanted to determine whether this interaction also impacts the bivalent domain marks H3K4me3 and H3K27me3. H3K4me3 is an activation mark and its occupancy on the promoters of some *Hox* genes is enhanced when expression of the *Hox* gene is increased, while at the same time, gene repressive mark H3K27me3 has decreased occupancy (Wang *et al*., 2011; Bertani *et al*., 2011). The reverse occurs when *Hox* gene expression is inhibited resulting in a lower occupancy of H3K4me3 and higher occupancy of H3K27me3. Chromatin immunoprecipitation (ChIP) were carried out as described in 4.2.16 (p. 90) and experiments showed that siRNA induced suppression of *mHotairm1* expression in NIH 3T3 cells (Fig. 5.20), resulted in a decreased occupancy of the activating mark H3K4me3 and an increased occupancy of repressive mark H3K27me3 in both *Hoxa1* and *Hoxa2* chromatin samples (Fig. 5.25 A-D). The occupancy of Histone 3 (H3) on both *Hoxa1* and *Hoxa2* gene chromatin samples was not affected by the down regulation of *mHotairm1* (Fig. 5.25 E, F). These results confirmed that *mHotairm1* does indeed regulate the expression of *Hoxa1* and *Hoxa2* via histone methylation impacting bivalent domain marks H3K4me3 and H3K27me3.

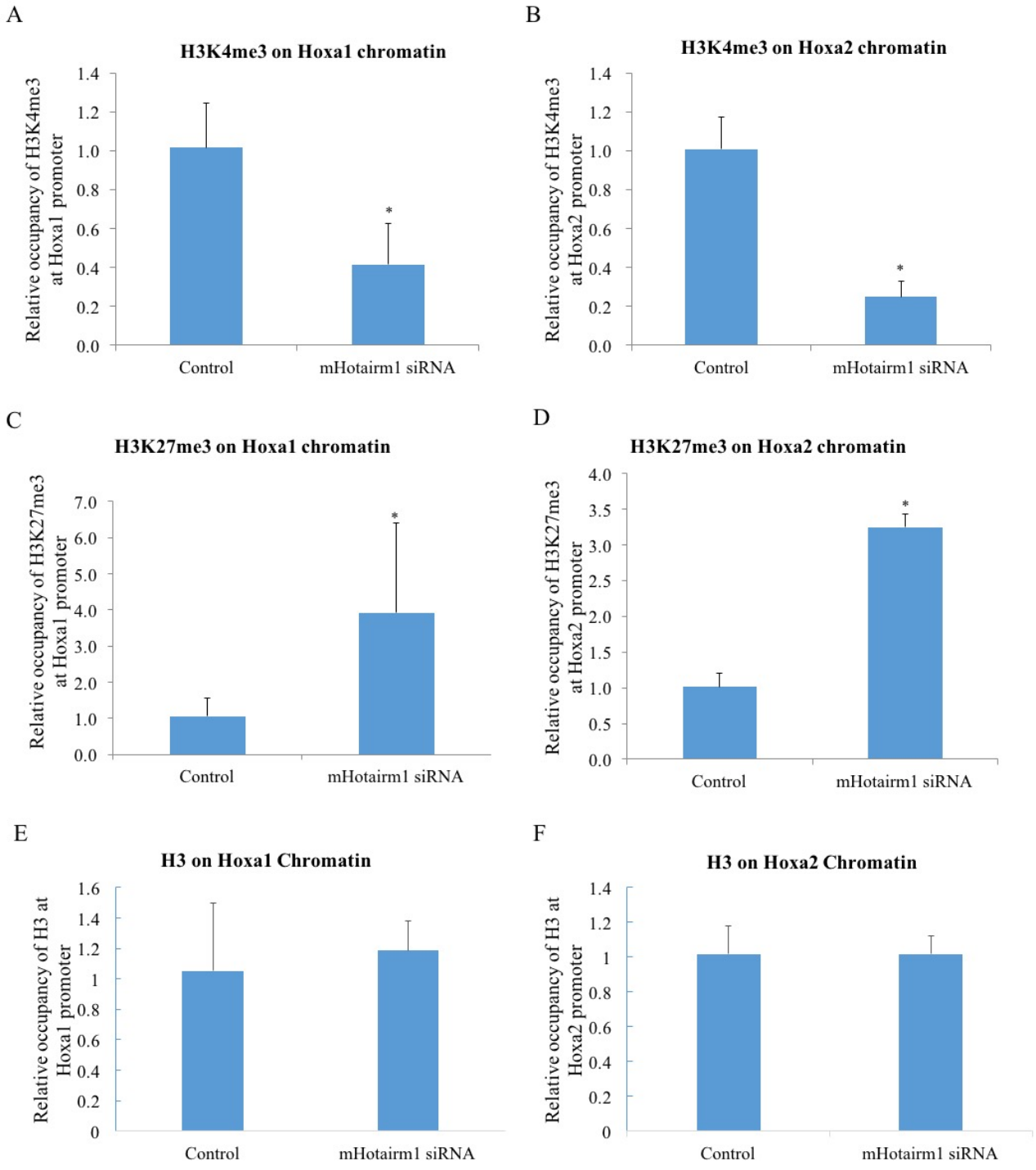


Figure 5.25. *mHotairm1* affects the bivalent histone methylation marks, H3K4me3 and H3K27me3, at *Hoxa1* and *Hoxa2* chromatin sites. qChIP was carried out as described in 4.2.16. Relative occupancy repreasts the fold enrichment of H3K4me3 (A,B), H3K27me3 (C,D) and H3 (E,F) on both *Hoxa1* and *Hoxa2* chromatin relative to input. Histone 3 was used as a positive control (E,F). Bars represent mean ±SEM, n =3. *p ≤ 0.05 compared to respective control.

The above findings demonstrate that *mHotairm1* can recruit histone methyltransferase complex MLL1/WDR5 to *Hoxa1* and *Hoxa2* gene and induce H3K4me3 to activate their expression. These findings are similar to the findings of *HOTTIP* (Wang *et al*., 2011) and *Mistral* (Bertani *et al*., 2011), which can positively regulate neighbouring *Hox* genes through the recruitment of MLL1/WDR5 complex.

5.3.4 WDR5 is sumoylated and this modification maybe important for the function of WDR5.

5.3.4.1 MLL1 specifically interacts with 50 kDa WDR5.

Using a WDR5 antibody in western blot, two distinct WDR5 bands at ~37KDa and at ~ 50KDa were observed in NIH 3T3 cell lysates (Fig. 5.22). In addition, GST-MLL1 (3810-3963) protein pull down of WDR5 confirmed a direct interaction between MLL1 and WDR5 but interestingly, only predominantly with the ~50 kDa isoform (Fig. 5.26). This ~50 kDa band was suspected to represent a modification of WDR5 and since only the ~50 kDa band co-precipitated with MLL1, this modification may be biologically important. The size difference between the ~50 kDa band and the unmodified WDR5 (~37 kDa) suggested the modification might be sumoylation, which is ~ 12 kDa in size (Kumar and Zhang, 2015). Indeed, in the presence of the ubiquitin/SUMO protector NEM, protein samples showed a stronger band intensity further supporting this hypothesis (Fig. 5.26A). Thus I hypothesized this post-translation modification to be sumoylation and additional experiments were carried out to test this hypothesis.

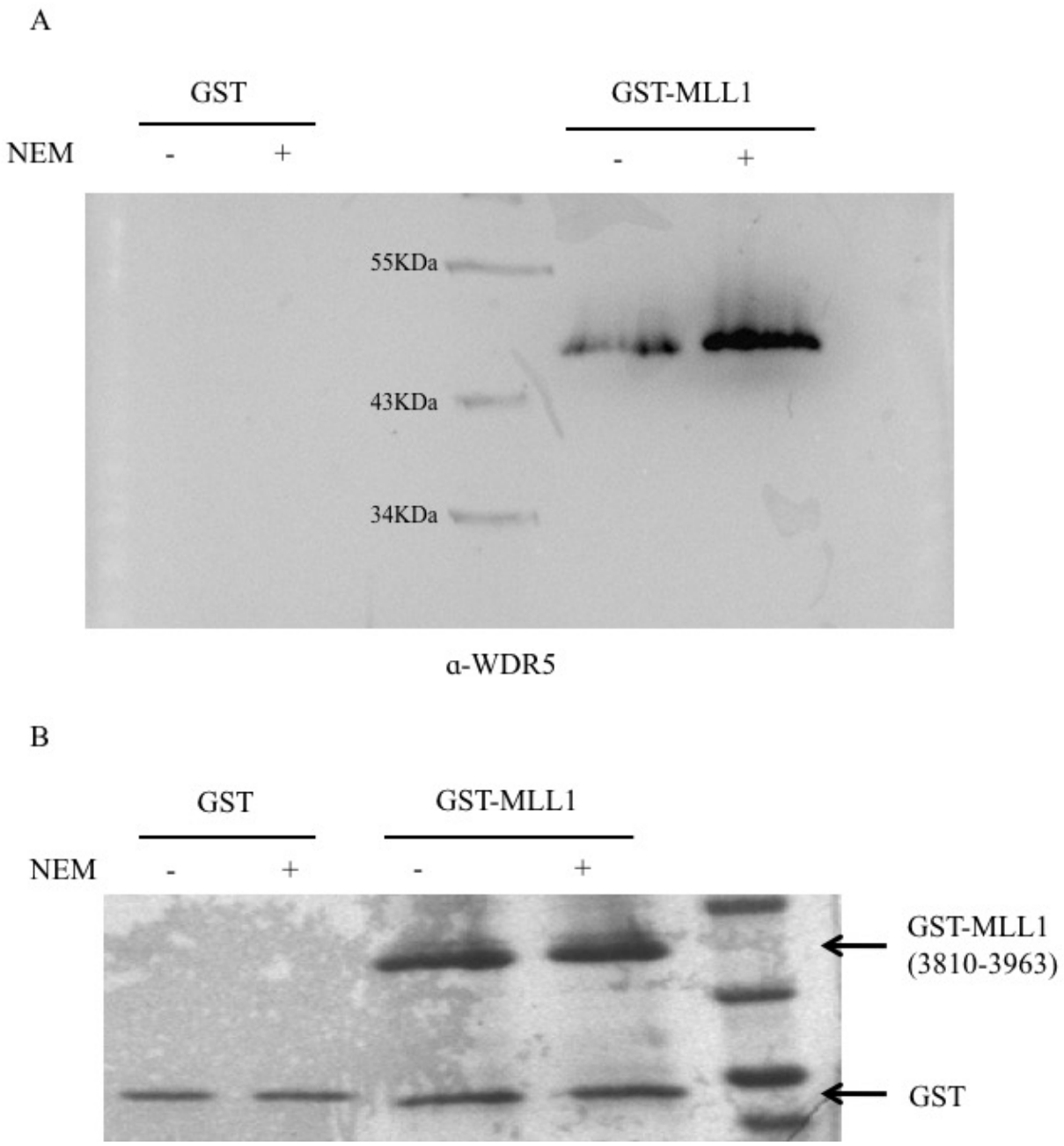


Figure 5.26. GST fused MLL1 (3810-3963) pull down of modified WDR5. A. Western blot showing modified WDR5 bands at ~ 50kDa in samples pulled down with GST fused MLL1 (3810-3963). The unmodified WDR5 band (~37kDa) was not observed. NEM: ubiquitin/SUMO protector N-ethylmaleimide. GST was used as a negative control. B. Loading control of GST and MLL1 (3810-3963). Coomassie blue staining showing equal loading of GST and GST fused MLL1 used in the pull down experiment above.

5.3.4.2 Sumoylated WDR5.

I carried out immunoprecipitation assays and established that the protein samples precipitated with WDR5 antibody could be detected with a SUMO1 antibody in Western blot assays (Fig. 5.27A). Similarly, protein samples precipitated with SUMO1 antibody could also be identified with the WDR5 antibody (Fig. 5.27B). Hence, the modification of WDR5 as it exists in NIH 3T3 cells is most likely sumoylation. Experiments using mass spectrometry to confirm sumoylation of WDR5 will be needed to characterize this further.

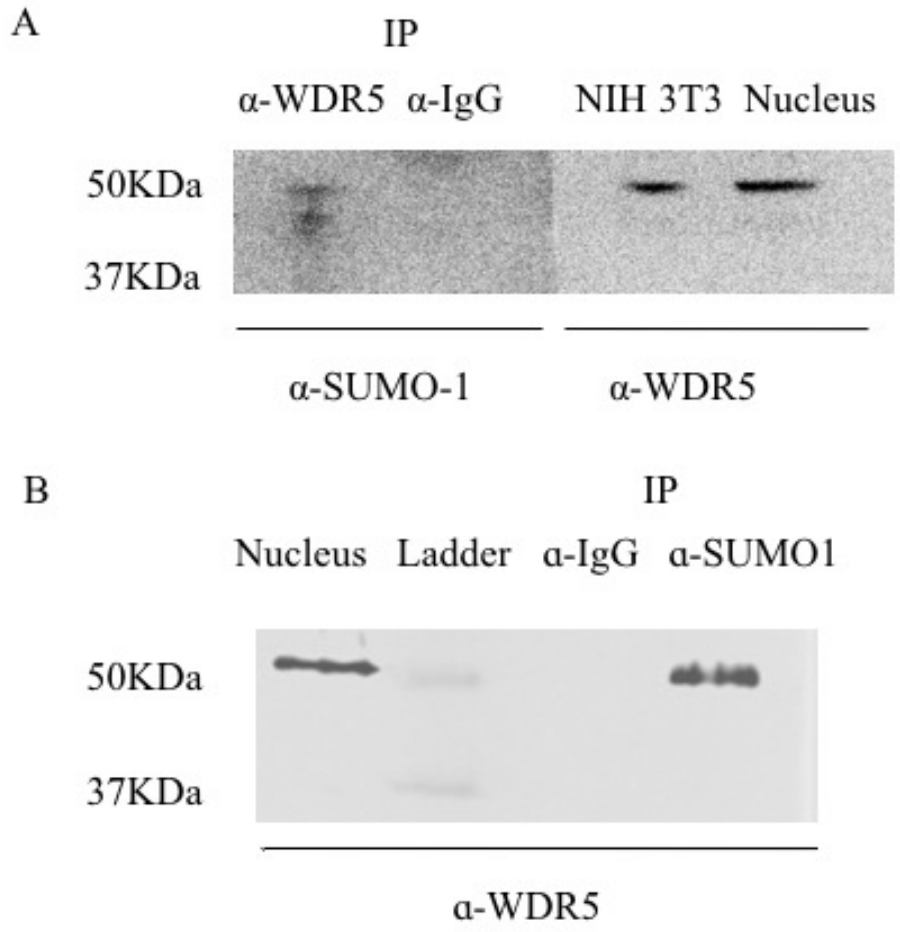


Figure 5.27. Reciprocal co-immunoprecipitation of WDR5 and SUMO1. A. SUMO1 antibody detected a band in protein samples immunoprecipitated with WDR5 antibody at the exact size as modified WDR5 seen in NIH 3T3 nucleus. B. WDR5 antibody also detected a band in protein samples immunoprecipitated with SUMO1 antibody at the size of modified WDR5. IgG antibody was used as the negative control.

5.3.4.3 Cellular distribution of WDR5 in NIH 3T3 cells.

Although SUMO modification of proteins has multiple functions, one important function is their role in nuclear-cytosolic transport (Eun Jeoung *et al*., 2008; Berndt *et al*., 2012; Lamoliatte *et al*., 2014). I investigated this further by first performing immunocytochemistry to determine the distribution profile of WDR5 in NIH 3T3 cells. Immunocytochemistry revealed that WDR5 is distributed throughout the cell in both cytoplasm and nuclei (Fig. 5.28). Subsequently, nuclei and cytoplasm were isolated from the NIH 3T3 cells using the method described in section 4.2.17. Western blot analysis showed that the modified WDR5 protein is present in the nuclei whereas both forms of WDR5 exist in the cytoplasm (Fig 5.29). The proteins β-actin (only expressed in cytoplasm) and Histone 3 (H3, only expressed in nucleus) were used to verify the separation of nuclei and cytoplasm.

We can speculate that since only the modified WDR5 is present in the nuclei, this SUMO modified WDR5 may play a role in translocation of the protein from the cytoplasm to the nuclei.

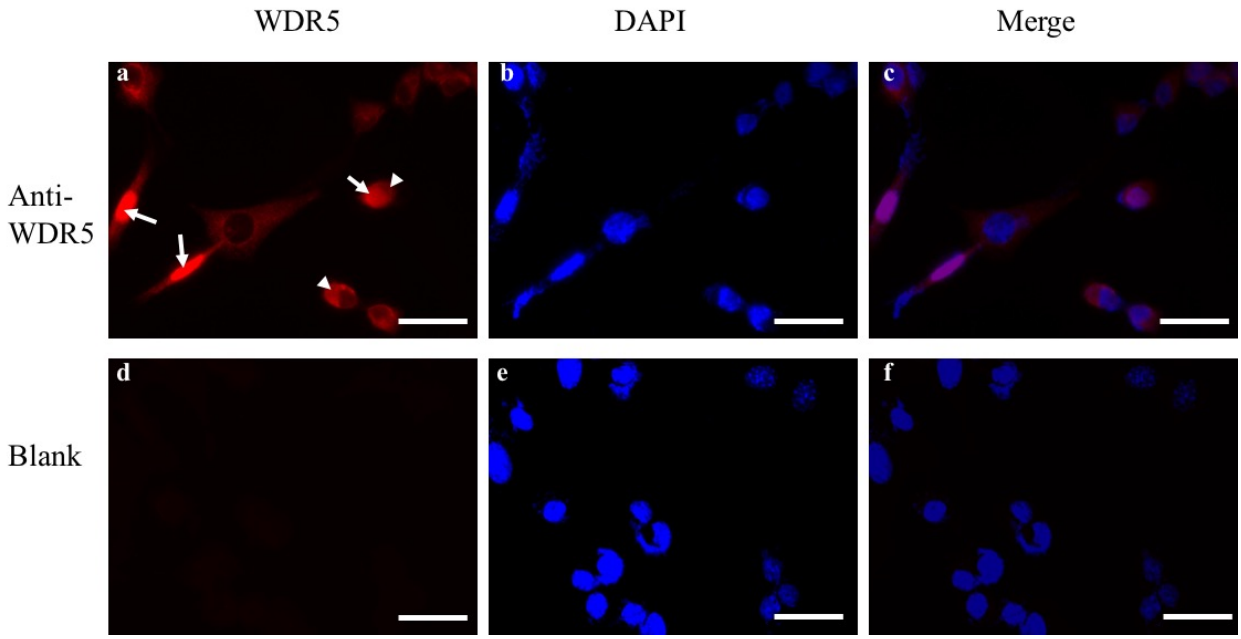


Figure 5.28. WDR5 is present in both cytoplasm and nuclei in NIH 3T3 cells. NIH 3T3 cells were fixed and immunostained with a WDR5 antibody (shown in red, a and c). Nuclei were labelled with DAPI (shown in blue, b-c, e-f). A blank group without primary antibody (WDR5 antibody) is used as negative control (d). Arrows point to nuclei and arrowheads point to cytoplasm. Scale bars = 3 microns.

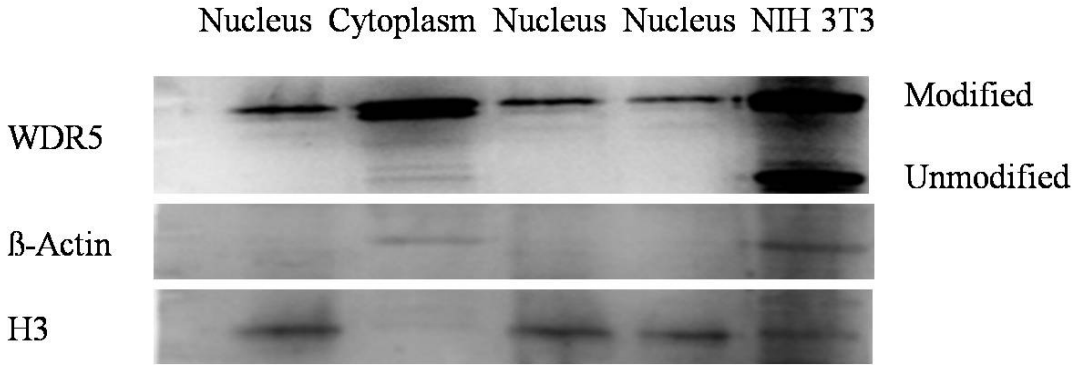


Figure 5.29. WDR5 proteins in NIH 3T3 cell cytoplasm and nuclei. Western blot analysis shows modified WDR5 protein band (~50kD) in nuclei, cytoplasm and NIH 3T3 whole cell lysates. Unmodified WDR5 protein band (~37kD) was only observed in the cytoplasm and NIH 3T3 cell lysates. β-Actin protein was observed only in cytoplasm and NIH 3T3 whole cell lysates. H3 protein was observed in nuclei and NIH 3T3 whole cell lysates.

5.3.4.4 Sumoylated WDR5 is responsive to ATRA induction.

Since ATRA induces the expression of *mHotairm1* as well as *Hoxa1* and *Hoxa2* (Fig. 5.19), I further investigated whether ATRA would have any effect on post-translation modification of WDR5. Interestingly, ATRA treatment appeared to show a stronger band intensity of the modified WDR5 (~50 kDa) protein in the nuclei of NIH 3T3 cells (Fig 5.30). As *mHotairm1* can positively regulate *Hoxa1* and *Hoxa2* expression through the regulation of H3K4me3 occupancy on their promoters, it is highly possible that exposure to ATRA which induces *Hoxa1* and *Hoxa2* expression is through the increase of both modified WDR5 and *mHotairm1* expression. In summary, my results reveal a new a gene regulatory pathway in which *mHotairm1* working closely with H3K4 methylation complex and sumoylated WDR5, regulates *Hoxa1* and *Hoxa2* expression.

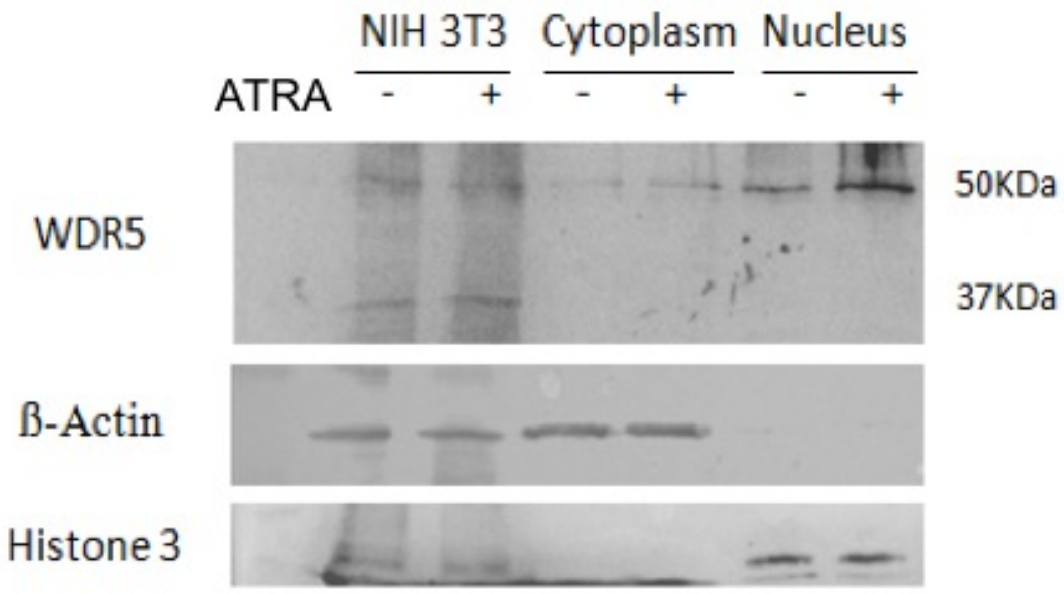


Figure 5.30. Effect of ATRA on the expression sumoylated WDR5. After ATRA treatment, an increased sumoylated WDR5 in NIH 3T3 nuclei was observed. β-actin bands were observed in cytoplasm and NIH 3T3 cell lysates. Histone 3 bands were observed in the nuclei and NIH 3T3 cell lysates.

These findings suggest that WDR5 is sumoylated and this modification plays an important role in its interaction with *mHotairm1* and MLL1. This modification may help WDR5 to translocate from cytoplasm to the nucleus. Further experiments are needed to study the function of sumoylated WDR5.

CHAPTER 6

6. Discussion

6.1 DNA methylation

DNA methylation is an important epigenetic mechanism in mammalian embryo development (Bird, 2002; Santos *et al*., 2005; Bartolomei and Ferguson-Smith, 2011; Lomvardas and Maniatis, 2016). DNA methylation which primarily occurs on cytosine bases in GpG rich sequences of the promoter regions leads to gene silencing through the inhibition of transcription factor binding and the changing of chromatin structure into a repressive state (Domcke *et al*., 2015). Our lab has been interested in *Hoxa2* regulation in mouse palatal development and has revealed a spatio-temporal expression pattern of *Hoxa2* during palatogenesis (Nazarali *et al*., 2000; Smith *et al*., 2009, 2013). It is not known whether this differential expression of *Hoxa2* is regulated by a specific epigenetic mechanism(s). The genome is comprised of many CpG rich regions and many are located near 5′ regulatory regulatory regions of genes. My analysis of *Hoxa2* promoter revealed three CpG islands close to *Hoxa2* 5′ regulatory regulatory region. I have used several approaches to study these CpG islands. First I used MSP to study the three CpG islands within the *Hoxa2* promoter region. The advantage of this method is that by designing methylated and unmethylated specific primers, the methylation status can be tested by a simple PCR. However, since PCR primers are usually only 18-24nt long, the CpG sites that can be covered by forward and reverse primers are limitied. This also resulted in higher degree of primer sequence similarity between methylated and unmethylated primers. A critical PCR condition is needed for MSP primers to selectively bind and amplify specific methylated/unmethylated target sequences. To develop optimal PCR conditions for all sets of

MSP primers, DNA samples from NIH 3T3 cells and EG7 cells were used as unmethylated and methylated control, respectively. It has previously been shown in our lab that the *Hoxa2* promoter is unmethylated in NIH 3T3 cells and that it remains highly methylated in the EG7 cells (Wang and Nazarali, unpublished). After testing several different annealing temperatures, I was able to develop optimal PCR cycle conditions for each pair of MSP primer set. In the developing mouse palate, DNA samples collected at all four stages of palatogenesis were recognized by unmethylated primers but not with methylated primers, suggesting *Hoxa2* promoter is most likely unmethylated during palate development.

The limitation of using MSP is that only CpG sites recognized by the primers can be tested. To examine the methylation status of each single CpG site, a BSP method was used. BPS is a method that uses PCR to amplify a certain CpG island, whether methylated or unmethylated, and the methylation status of each individual CpG site can then be analyzed by DNA sequencing. This method requires a longer experimental process but the advantage is that each GpG site within the CpG island can be analyzed. Since the CpG region-1 is closest to *Hoxa2* transcription start site and likely most relevant to *Hoxa2* expression, this region was selected for further BSP analysis. The results of BSP showed that all 14 CpG sites tested in the CpG region-region 1 were unmethylated in NIH 3T3 cells and in developing mouse palate from E12 to E15, whereas 10 of the 14 CpG sites were found to be methylated in EG7 cells. These results supported my findings from the MSP experiments. From these observations we can conclude that during mouse palate development the *Hoxa2* promoter within the CpG-region 1 primarly remains unmethylated and that the DNA methylation status of the *Hoxa2* promoter does not change with the spatio-temporal expression of *Hoxa2* during palatogenesis.

6.2 Regulation of *Hoxa2* gene expression by miRNAs

Currently, there are several studies that focus on miRNA regulation of gene expression (Yan and Jiao, 2016; Green *et al*., 2015; He *et al*., 2015; Usmani *et al*., 2016). Regulation of gene expression by miRNAs is complex because a specfic gene can be regulated by many different miRNAs and whereas a specific miRNA can also regulate expression of several different genes (Doench and Sharp, 2004; Felekkis *et al*., 2010). Generally, miRNAs bind to the 3'UTR of a gene, hence I performed in-silico analysis and identified six miRNAs that had the potential to bind 3'UTR of the *Hoxa2* gene. Following this, I investigated expression profiles of these miRNAs in mouse palatal tissues, and in cell lines in which *Hoxa2* expression is already studied (Smith *et al*., 2009; Wang and Nazarali unpublished). Since these predicted miRNAs were considered to downregulate *Hoxa2* expression, I primarily focused on miRNA expression patterns that were complementary to that of the *Hoxa2* expression. The two miRNAs: miR-431 and miR-298 showed high expressions in NIH 3T3 cells at the same time as *Hoxa2* expression. Thus these two miRNAs are unlikely to down regulate *Hoxa2* in NIH 3T3 cells. All six miRNAs exhibited low expression levels in EG7 cells, possibly due to globally repressed transcription activity and high DNA methylation status of gene promoters in EG7 cells (Wang and Nazarali unpublished). I then investigated the expression of of all six miRNAs in developing mouse palate at stages E12 to E15. Three miRNAs (miR-669b, miR-376c and miR-431) exhibited increased expression from E13 to E15 in mouse palate, during the same period when *Hoxa2* expression is declining (Smith *et al*., 2009). Since miR-431 showed high expression in NIH 3T3 cells, I continued further investigations with only miR-669b and miR-376c.

To investigate whether miR-669b and miR-376c are able to down regulate the expression of *Hoxa2*, I transfected synthesized mature miRNAs: miR-669b mimic and miR-376c mimic into

NIH 3T3 cell cultures. Overexpression of both miRNAs by miRNA mimics resulted in the downregulation of *Hoxa2* expression at both transcriptional and translational level. At this stage, it was not known whether miR-669b and miR-376c miRNAs bound directly to the 3′UTR of *Hoxa2* or had an indirect effect on *Hoxa2* expression via regulation of upstream genes that may control *Hoxa2* expression. A luciferase expressing vector carrying the 3′UTR of *Hoxa2* was used in luciferase assays to determine whether miR-669b and miR-376c miRNAs bound to the *Hoxa2* 3′UTR. Luciferase assays showed that miR-669b appears to directly bind to *Hoxa2* 3'UTR since mutations within the seed sequences abrogated luciferase activity, however miR-376c had no effect on luciferase assay and so its effect on *Hoxa2* expression may likely be via an indirect effect. Although most miRNAs bind to the 3'UTR of their target genes, some miRNAs are also known to bind to coding regions of mRNA (Hu *et al*., 2011). Hence, it is possible that miR-376c may affect *Hoxa2* expression by interacting with other regions of the *Hoxa2* mRNA. An additional possibility is that miR-376c may be regulating genes upstream of *Hoxa2* which in turn is regulating *Hoxa2* expression. One such possibility is the Krox20 gene that is upsteam of *Hoxa2* (Nonchev *et al*., 1996). In human pluripotent stem cells, miR-376c has been reported to inhibit the expression of SMAD4 gene in the TGF-β (transforming growth factor β) pathway (Liu *et al*., 2014). SMAD4 is known to bind directly to PAX6 (paired box 6) gene promoter and suppress its expression (Liu *et al*., 2014). Thus, overexpression of miR-376c would lead to the downregulation of SMAD4 expression which inturn would lead to an upregulation of PAX6 expression. Pax6 has been reported to induce the expression of Nab1 (NGFI-A Binding Protein 1), a Krox20 repressor in the chick embryo (Kayam *et al*., 2013). Hence, overexpression of miR-376c could lead to an upregulation Nab1 expression via Pax6 and subsequent downregulation of Krox20 expression. Krox20 is a known upstream gene of *Hoxa2* and can directly activate its

expression (Nonchev *et al*., 1996). Thus, downregulation of Krox20 would induce downregulation of *Hoxa2* expression. Hence, an indirect link between miR-376c and *Hoxa2* gene expression may exist via genes upstream of *Hoxa2* (Fig 6.1).

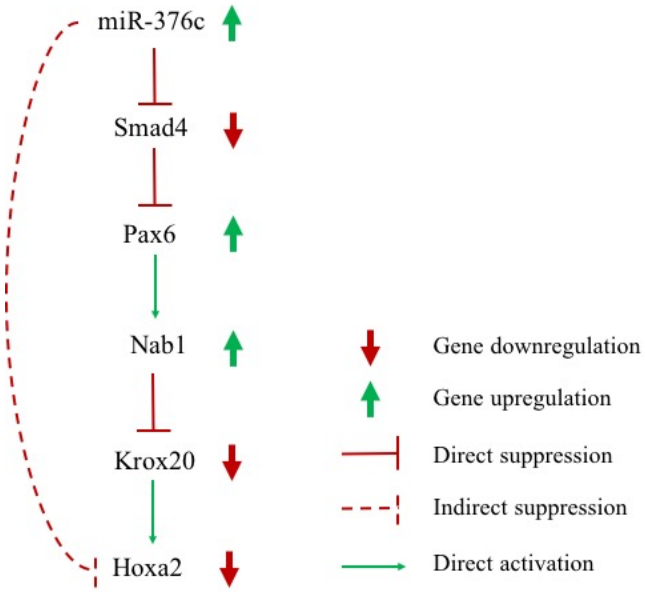


Figure 6.1. A putative indirect link between miR-376c and *Hoxa2* expression. miR-376c can directly inhibit the expression of Smad4 which in turn activates Pax6 expression. Pax6 induces the expression of Nab1 which in turn inhibits the expression of Krox20, an upstream *Hoxa2* activator.

Researchers in the past have generally put a higher emphasis on comparing differences within coding regions of genes between various species. However, many researchers are beginning to realize that the non-coding regions have important regulatory roles and sequence differences in these regions could lead to differentially regulated gene expression between various species (Barrett *et al*., 2012). There are two predicted miR-669b binding sites on mouse *Hoxa2* 3'UTR. My mutation experiments showed that the two binding sites appear to work independently from each other. I analyzed the degree of sequence similarity for both miR-669b binding sites and found that the binding site 1 is evolutionally conserved among the five species

examined, human, mouse, rat, chimpanzee and dog. However, the binding site 2 in mouse only exibited sequence similarity with that in rat and not with the other three species examned (Fig. 5.5). Since the binding site 1 is present in all five species examined, it is possible that miR-669b has more conserved role in regulating *Hoxa2* gene expression via this binding site. While it is known that one miRNA may have hundreds of gene targets, having a single miRNA binding site may not be sufficient to strongly inhibit gene expression (Stark *et al*., 2005). Thus, it is also possible that the regulation of *Hoxa2* by miR-669b would not be as strong in other species compared to mouse or rat which has two miR-669b binding sites on *Hoxa2* 3′UTR.

Interestingly, evidences support genomic evolution of 3'UTR region to mitigate and augment the effects of miRNA on gene expression regulation (Stark *et al*., 2005; Thomsen *et al*., 2010). Some mRNAs have evolved to have different lengths of 3'UTRs so that they have different 'visibilities' to miRNAs in different tissues (Stark *et al*., 2005; Thomsen *et al*., 2010). Thomsen and colleagues found different 3' UTR splicing exist in several *Hox* genes in *Drosophila* (Thomsen *et al*., 2010). For example, the longer form of *Ubx* mRNA has eight miR-iab4/8 binding sites and is expressed mainly in the developing CNS. Expression levels of the longer *Ubx* mRNA with multiple miR-iab4/8 binding sites is very sensitive to miR-iab4/8 regulations. The shorter form of *Ubx* mRNA with only three miR-iab4/8 binding sites in the 3'UTR is mainly expressed in germ band elongation and removal of miR-iab4/8 binding sequence did not have a significant effect on the expression of the short form of the *Ubx* transcript (Thomsen *et al*., 2010). Differential splicing of *Hoxa2* 3'UTR has not been reported; however, bioinformatic analysis has revealed high frequency of alternative polyadenylation in mammals (Brett *et al*., 2000; Di Giammartino *et al*., 2011). Hence, it is still possible that *Hoxa2* may also have differentially spliced 3'UTR. Having a single miR-669b binding site in human

HOXA2 3'UTR may make it less 'visible' to miR-669b, although my experiments showed that each single miR-669b binding site on *Hoxa2* 3'UTR is sufficient to confer the repression of *Hoxa2* expression when miR-669b was over expressed. Possessing two binding sites may enhance the regulatory role of miR-669b on *Hoxa2* expression under normal biological conditions.

In my study, I have also found that in the developing mouse palate (from E13 to E15), miR-669b has a complementary expression to that of *Hoxa2*. At the E12 stage, both *Hoxa2* and miR-669b have relatively low expression in palate indicating that miR-669b may not be responsible for regulating *Hoxa2* expression at this stage in mouse palate. However, it is highly probable that miR-669b may play a role in down regulating *Hoxa2* expression towards the end of mouse palate development when *Hoxa2* palatal expression declines and miR-669b expression is the highest. Since *Hoxa2* can also regulate ear and nervous system development (Minoux *et al*., 2013; Cox *et al*., 2014; Gavalas *et al*., 1997; Wang *et al*., 2011), it may be of interest to investigate whether miR-669b plays a role in these developmental processes. Many pathological processes, including cancer, neurological diseases and cardiovascular diseases impact specific miRNA expression changes, making them a promising biomarker and potential diagnostic tool for a variety of diseases (Chi and Zhou, 2016; Stoicea *et al*., 2016; Wallace *et al*., 2016). The regulatory roles of miRNAs also render them promising therapeutic targets in many diseases (Broderick and Zamore, 2011). The mechanisms by which miRNAs regulate gene expression are still not fully understood and further studies on miRNAs will lead to a better understanding of biological processes.

In summary, my investigations identified two miRNAs, miR-669b and miR-376c, that regulate *Hoxa2* expression in mouse NIH 3T3 cells. MiR-669b directly binds to mouse *Hoxa2*

3'UTR to down regulate its expression, also two functional miR-669b binding sites were identified on *Hoxa2* 3'UTR. No direct interactions between miR-376c and *Hoxa2* 3'UTR was identified. miR-376c possibly binds to other regions of *Hoxa2* mRNA or other regulatory genes upstream to *Hoxa2* which in turn may regulate *Hoxa2* expression (Fig. 6.1). miR-669b is expressed in developing mouse palate with a complemenatry expression profile to *Hoxa2* and may have regulatory role in mouse palate development.

6.3 lncRNA *mHotairm1* regulation of *Hoxa1* and *Hoxa2* expression

Following the discovery of the lncRNA:Polycomb repressive complex 2 (PRC2) interaction (Zhao *et al*., 2008), significant interest has been generated to characterize the role of lncRNAs in epigenetic activation/inactivation of gene expression (Cao, 2014). Additionally, investigation on the transcriptional activity of the human *HOX* loci showed that many of the intergenic regions are actively transcribed and most of transcripts from the intergenic regions are lncRNAs (Rinn *et al*., 2007). *Hox* lncRNAs likely play important biological roles based on the fact that: 1) Some lncRNAs are conserved during evolution; 2) Like *HOX* genes, lncRNAs in HOX cluster have different expression patterns along the A-P axes depending on their physical location on the chromosome (Rinn *et al*., 2007). However, very few lncRNAs from *Hox* clusters have been chacterized. In my study, I have found a new lncRNA located between mouse *Hoxa1* and *Hoxa2* and further demonstrated its ability to recruit MLL1/WDR5 to nearby target genes and to regulate these gene expressions. My results provide additional evidence to support the hypothesis raised by Rinn *et al*. that transcription of lncRNA in *cis* may recruit TrxG proteins such as MLL1 and WDR5 to chromatin, leading to H3K4me3 induction and gene activation (Rinn *et al*., 2007).

6.3.1 Identification of *mHotairm1*

The existence of *HOTAIRM1* in human was reported by Zhang and colleagues in 2009 (Zhang *et al*., 2009). In GenBank database that I searched, a similar long non-coding transcript was found to potentially exist in the mouse. I made several attempts to identify the full length of *mHotairm1*. First, based on the predicted *mHotairm1* sequence provided by GenBank, I designed three pairs of PCR primers that covered 462bp of the predicted *mHotairm1* sequence in total. Each pair of primer set were designed to have overlapping region with the next pair. All three pairs of primers successfully amplified DNA fragments from cDNA samples generated from total RNA isolated from NIH 3T3 cells. Each amplified fragment was sequenced and matched exactly with the predicted sequence (Fig. 5.16B, Appendix 7-9). The analysis of the predicted *mHotairm1* sequence showed that this transcript has two exons and splicing occurs between nucleotide 162 and 165. Primer set Mush-138 covered a length from nucleotide 131 to 268 (Table 4.3, Appendix 8) which include the splicing site. The fact that primer set Mush-138 is only able to amplify DNA fragment from cDNA sample but not from genomic DNA (Appendix 12) further supported the existence of *mHotairm1* and ruled out a possible contamination from genomic DNA. This primer set was also used to identify *mHotairm1* in further experiments.

I then tried to identify the 3' and 5' end of *mHotairm1* using Rapid Amplification of cDNA Ends (RACE) approach. The strategy of RACE experiment is to add special adaptor sequences to both 3' and 5' end of mRNAs and construct a cDNA library. Then by using gene specific primers and adaptor primers, target gene fragments can be amplified with their 3' and 5' ends. Unfortunately, I was not able to successfully amplify either 3' or 5' end of *mHotairm1*. I have tried different *mHotairm1* specific primers but none of amplification products turned out to be from *mHotairm1*. The possible reasons are: in RACE experiments, only one gene specific

primer is used in target fragment amplification, which reduced the specificity of amplification compared to normal PCR, and secondly, lncRNAs usually have significantly lower levels of expression compared to protein-coding genes (Cabili *et al*., 2011; Derrien *et al*., 2012), which makes them even more difficult to amplify in a RACE experiment. However, I was able to identify the majority of the *mHotairm1* sequence (462nt) and confirm the existence of a new long noncoding transcript located between *Hoxa1* and *Hoxa2* in mouse.

6.3.2 *mHotairm1* expression in cell lines and mouse tissues

Most lncRNAs have tissue specific expression (Derrien *et al*., 2012). *mHotairm1* expression was identified in different mouse embryonic tissues (Fig 5.17). In the developing palate (E14) *mHotairm1* expression is localized to the medial edge epithelia (MEE) (Fig 5.18A). In the developing mouse palate at E14 the two vertical palate shelves begin elevating above the tongue and grow towards each other. The two palate shelves eventually contact each other at the MEE. The MEE layers from the two palate shelves then merge together to form the medial edge seam (MES) and finally fuse together (Ferguson, 1988). This process is regulated by a series of cellular and biochemical reactions and is a very important step in palate development (Ferguson, 1988; Jin and Ding, 2006, Smith *et al*., 2013, Lan *et al*., 2015). *In situ* hybridization histochemistry of E14 developing palate showed the expression of *mHotairm1* in MEE (Fig. 5.18), indicating *mHotairm1* may play a role in the palatal fusion. *mHotairm1* expression in E15 palatal shelves was only faintly detected by *in situ* hybridization histochemistry, which agrees with the real-time PCR data where *mHorairm1* expression has significantly declined. The expression of *mHotairm1* in E12 and E13 mouse palatal shelves was not detectable with *in situ* probes, possibly due to a low abundance of lncRNAs at these stages.

6.3.3 Regulation of *Hoxa1* and *Hoxa2* expression by *mHotairm1*

Although the presence of non-coding transcripts was initially considered to be "noise" and of little consequence, they have gained importance as more information is gathered on their important regulatory roles (Rinn and Chang, 2012; Bonasio and Shiekhattar, 2014). As Rinn and colleagues (2007) have suggested, the lncRNAs transcribed from Hox loci can either up or down regulate target gene expression. Thus to study the function of *mHotairm1*, I first tested whether a change in the expression of *mHotairm1* could influence the expression of nearby genes such as *Hoxa1* and *Hoxa2*. I chose to down regulate expression of *mHotairm1* with siRNA in mouse fibroblast cell line NIH 3T3 where *Hoxa1*, *Hoxa2*, as well as *mHotairm1* are all expressed. My results showed that downregulation of *mHotairm1* resulted in significant downregulation of *Hoxa1* and *Hoxa2* expression.

A well characterized activator of *Hox* gene expression, all-*trans* retinoic acid (ATRA) has also been shown to induce the expression of Hox loci derived lncRNAs (Zhang *et al*., 2009; Bertani *et al*., 2011). Moreover, the induction of 3′ HOX gene expression by ATRA is attenuated if the regulatory lncRNA *HOTAIRM1* is silenced (Zhang *et al*., 2009). As well Bertani and colleagues (2011) have reported that the transcription of lncRNA *Mistral* precedes its target gene *Hoxa6* and *Hoxa7* expression following ATRA treatment. These evidences support induction of *Hox* genes via the effect of ATRA on lncRNA expression. My results showed that ATRA induces the expression of *mHotairm1*, *Hoxa1* and *Hoxa2* in NIH 3T3 cells. Based on the findings of Bertani *et al* (2011) and Zhang *et al* (2009), it is highly probable that ATRA induces the expression of *mHotairm1* further leads to an increase in the expression of *Hoxa1* and *Hoxa2*. Thus my findings indicate that *Hoxa1* and *Hoxa2* have the same expression trend as that of *mHotairm1* and that *mHotairm1* can positively regulate the expression of *Hoxa1* and *Hoxa2*.

6.3.4 *mHotairm1* activates *Hoxa1* and *Hoxa2* expression through histone methylation

LncRNAs are involved in numerous cellular processes, including ES cell pluripotency, cellular pathway regulation and cell-cycle regulation (Rinn and Chang, 2012). They are also important during development (Schmitz *et al*., 2016) and in certain diseases, such as cardiovascular disease (Archer *et al*., 2015), nervous system diseases (Briggs *et al*., 2015) and cancer (Schmitt and Chang, 2016). Despite all the different cellular processes and diseases that involve lncRNAs, one fundamental rule has emerged: lncRNAs drive the formation of ribonucleic-protein complexes to regulate gene expression (Rinn and Chang, 2012). One large group of protein partners that bind to lncRNAs are the chromatin and DNA modification complexes. These complexes include DNA methylation complex, PRC1 and PRC2, TrxG complex and HDACs (reviewed by Rinn and Chang, 2012), which closely link lncRNA with other epigenetic marks. The lncRNAs transcribed from *Hox* loci have thus far been reported to bind to two histone methylation complexes: PRC2 which tri-methylates H3K27 and suppresses gene expression (Rinn *et al*. 2007), and MLL1/WDR5 complex, which belongs to TrxG and tri-methylates H3K4 to activate gene expression (Wang *et al*., 2011; Bertani *et al*., 2011). Since my results show that *mHotairm1* plays a role in activating the expression of *Hoxa1* and *Hoxa2*, I speculated that *mHotairm1* may achieve its gene activation role by recruiting MLL1/WDR5 complex to the chromatin of *Hoxa1* and *Hoxa2*.

Capture hybridization analysis of RNA targets (CHART) developed by Simon *et al* (2011) is method used to enrich endogenous RNAs along with their associated proteins and their DNA targets. Similar to ChIP assays, samples are cross-linked and chromatin is sheared into small fragments. Then instead of using antibodies targeting chromatin proteins to enrich target fragments as in ChIP, affinity-tagged oligonucleotides are used to enrich target lncRNA together

with its protein partners and chromatin targets. In this way, CHART can achieve two goals at the same time: to study the genomic binding sites of a lncRNA and to study which protein complex is associated with this lncRNA. The advantages of CHART over ChIP in the study of lncRNA function are: (1) CHART can allow for the study of protein, DNA and RNA from the same coprecipitated sample; (2) CHART is able to identify the genomic target of lncRNA. The key step in CHART experiment is to design a lncRNA specific oligo probe to specifically bind to and enrich the target RNA. Since lncRNAs have secondary structures and are usually bound to proteins, the regions on the lncRNA that can be recognized by a complimentary oligo probe are limited. Additionally, the oligo probe still should be long enough to specifically recognize target lncRNA. Thus, I designed three 25-mer *mHotairm1* specific probes targeting 3′ (Probe 1), middle (Probe 2) and 5′ (Probe 3) part of *mhotairm1* (Fig. 4.4). The samples enriched with three *mHotairm1* specific probes were compared with samples pulled down with a control probe (Probe C) to rule out any background interference and nonspecific binding. Probe 2 showed the strongest binding to *mHotairm1*. The DNA samples retrieved from Probe 2 pull down showed that *Hoxa1* and *Hoxa2* as well as additional 3' HoxA genes are all targets of *mHotairm1* while the 5' HoxA gene *Hoxa13* did not specifically interact with *mHotairm1*. This result is supported by Wang *et al*'s (2011) finding that in HoxA cluster, 3' lncRNAs primarily interact with 3' HoxA genes while 5' lncRNAs primarily interact with 5' HoxA genes. The protein samples retrieved from the same pull down sample showed that the TrxG protein WDR5 is associated with *mHotairm1*, suggesting MLL1/WDR5 complex may play a role in the regulation of *Hoxa1* and *Hoxa2* gene expression by *mHotairm1*. Since histone acetylation is also a gene activation marker, I examined the presence of histone acetyltransferase CBP and found no interaction between *mHotairm1* and CBP. This, however, does not rule out the possibility that histone acetylation is

involved in *mHotairm1* induced expression of *Hoxa1* and *Hoxa2*, since there are other histone acetyltransferases that may be able to interact with WDR5 (Zhao *et al*., 2013; Dias *et al*., 2014). As CHART probes were DNA probes designed complementary to the sequence of *mHotairm1*, they may also have the abilities to bind to the genomic DNA, especially in region where *mHotairm1* is transcribed. To reduce the background due to the direct binding of genomic DNA to CHART probes, DNA samples can be eluted with RNase-H instead of biotin. In this way, only the DNA bound through RNA should be eluted from the resin, but not the DNA that is directly bound to the oligo probe (Simon *et al*., 2011).

WDR5 is a subunit of MLL1/WDR5 complex and is also involved in other functional complexes (Zhao *et al*., 2013; Dias *et al*., 2014). Thus, to confirm the involvement of histone methylation in the regulation of *Hoxa1* and *Hoxa2* expression by *mHotairm1*, I wanted to determine whether there is an interaction between *mHotairm1* and MLL1. Unfortunately, I was not able to find a good antibody that could recognize MLL1 in western blots, partially because MLL1 is a very large and complex protein. Thus, I was not able to determine the interaction between MLL1 and *mHotairm1* in CHART experiment. However, I designed a pull down experiment that would confirm the interaction between *mHotairm1* and MLL1/WDR5 complex. A GST tagged WDR5 (full length) and MLL1 (amino acid 3810-3963) expression vectors were constructed and a prokaryotic system (*E.coli*) was used to produce WDR5 and MLL1 proteins. As the mammalian MLL1 cDNA is ~12 kb in length and MLL1 protein contains ~4000 amino acids (Zhang *et al*., 2013), it is almost impossible to produce a full length MLL1 protein in *E.coli*. Bertani and colleagues reported that amino acids 3810 to 3963 is the critical region of MLL1 that binds lncRNA *Mistral* (Bertani *et al*. 2011), thus I selected this region of MLL1 in the pull down experiment. GST tag was used for the purification of proteins produced in *E.coli* and purified

WDR5 and MLL1 (3810-3963) proteins were applied to NIH3T3 whole cell lysate. Both WDR5 and MLL1(3810-3963) could pull down *mHotairm1*, which indicates that *mHotairm1* is closely associated with histone methytransferase complex MLL1/WDR5. However, it is not clear which of the two proteins bind directly to *mHotairm1*, since the fragment of MLL1 used in pull down experiments contains the Win motif, which is the region on MLL1 that also binds to WDR5 (Zhang *et al*., 2013). Hence possibility exists that the MLL1 pulled down *mHotairm1* is due to its interaction with WDR5.

There is also some debate in the literature as to which specific protein in the TrxG complex is key for the interaction with lncRNA. Wang and colleagues reported that lncRNA *HOTTIP* can only bind to WDR5 but not MLL1 (Wang *et al*., 2011). This is supported by Yang and colleagues who have identified several critical amino acids of WDR5 that are key in the binding of lncRNA *HOTTIP* (Yang *et al*., 2014). However, Bertani *et al.* (2011) demonstrated that *Mistral* binds directly to MLL1 only and not to WDR5. Nevertheless, my findings confirmed that *mHotairm1* interacts with MLL1/WDR5 complex to target *Hoxa1* and *Hoxa2* genes.

To further support the above findings that *mHotairm1* can recruit MLL1/WDR5 complex to *Hoxa1* and *Hoxa2* genes and regulate their expression, I downregulated the expression of *mHotairm1* in NIH 3T3 cells and investigated its impact on H3K4me3 and H3K27me3's occupancy on *Hoxa1* and *Hoxa2* chromatin using ChIP. The knockdown of *mHotairm1* resulted in reduced occupancy of gene activation mark H3K4me3 and increased occupancy of gene suppression mark H3K27me3, which correlates with my previous finding that downregulation of *mHotairm1* leads to decreased *Hoxa1* and *Hoxa2* expression.

These findings together suggested that *mHotairm1* can regulate *Hoxa1* and *Hoxa2* expression through the recruitment of histone methytransferase complex MLL1/WDR5. Hence,

the regulatory mechanism of *mHotairm1* appears to be very similar to *Mistral* (Bertani *et al*. 2011) and *HOTTIP* (Wang *et al*., 2011), other two lncRNAs identified within the HoxA cluster.

6.3.5 Modification of WDR5

Another significant aspect of my work is the finding that WDR5 is sumoylated and this modification is important for its function. Several pieces of information led me to suggest that WDR5 is being sumoylated. Firstly, the molecular mass difference between the modified and unmodified WDR5 is ~13KDa, which is close to the molecular mass of sumoylation (~11KDa) (Hay, 2005; Kumar and Zhang, 2015). Secondly, the 50KDa protein identified in my experiments is protected by NEM, a ubiquitin/SUMO protector (Fig. 5.26A). To confirm that WDR5 is sumoylated, I first used immunoprecipitation (IP) to test whether proteins immunoprecipitated by WDR5 antibody can be recognized by SUMO1 antibody, and proteins IP by SUMO1 antibody can be recognized by WDR5 antibody. In WDR5 antibody precipitated sample, a 50KDa band was detected by SUMO1 antibody, which is at the same molecular weight as that of the modified WDR5 in NIH 3T3 cells. To avoid the influence by IgG heavy chain, which is also at 50KDa, I used a rabbit WDR5 antibody (IgG) for IP and mouse SUMO1 antibody (IgG) for western blot. The sample precipitated with SUMO1 antibody can also be detected by WDR5 antibody at 50KDa. Again to avoid the influence of IgG heavy chain, a mouse SUMO1 antibody (IgG) was used for IP and a mouse WDR5 antibody (IgM) was used in western blot. The IP experiment supported the hypothesis that WDR5 is being sumoylated. This finding is supported by Nayak *et al.*'s work demonstrating that WDR5 is sumoylated in an *in vitro* transcription/translation system (Nayak *et al*., 2014). I further tried to use mass spectrometry to confirm my finding and to determine which amino acid is sumoylated to better

study the function of sumoylated WDR5. I enriched the modified WDR5 using immunoprecipitation and ran the protein sample on a SDS-PAGE gel that was subsequently stained with coomassie blue. Unfortunately, the IgG heavy chain from the WDR5 antibody also showed up at 50KDa in the coomassie blue stained gel, thus I was not able to distinguish the modified WDR5 from IgG.

Even though I could not conclusively determine which amino acid in WDR5 is sumoylated, I was able to show the importance of this modification from my studies. From the CHART experiment, only modified WDR5 (~50KDa) was observed in samples pulled down with *mHotiarm1* specific probes, even when two WDR5 protein (modified and unmodified) bands were detected in NIH 3T3 cell lysates. In addition, when I used GST-fused MLL1 (amino acid 3810-3963) to pull down protein samples from NIH 3T3 cell lysates, again only the modified WDR5 band was observed in western blot assays. Immunocytochemistry showed that WDR5 proteins were present in both cytoplasm and nucleus of NIH 3T3 cells. I then analyzed the distribution of modified and unmodified WDR5 in NIH 3T3 cells. Results indicated that there is only modified WDR5 in the nucleus while both forms exist in the cytoplasm. This result further supported the importance of the modification of WDR5 and indicated that this modification may facilitate WDR5 translocation to the nucleus. Sumoylation have multiple functions for proteins and translocation of proteins to nucleus is one of its key functions (Kumar and Zhang, 2015), which further supports my findings. During the separation of cytoplasm and nucleus, only weak unmodified protein band of WDR5 could be observed in cytoplasm (Fig. 5.28, Fig. 5.29), which is possibly due to the dilution of sample during the separation. Because for the separation method I used, protein samples in the cytoplasm are much more diluted compared to nucleus samples and

may be the reason why I observed a lower concentration of the unmodified WDR5 in the cytoplasm.

Treatment with ATRA (all-*trans* retinoic acid) also impacted sumoylated WDR5, appearing to increase the quantity of sumoylated WDR5 in the cell nucleus but not influence the total quantity of WDR5 protein in NIH 3T3 cells. These results indicate that ATRA is able to influence the distribution of modified WDR5 in NIH 3T3 cells and most likely facilitates its translocation to the nucleus. Together, all these findings suggested that sumoylated WDR5 with its interaction with *mHotiarm1* plays an important role in impacting changes in H3K4me3 and H3K27me3 occupancy in epigenetic regulation of *Hoxa1* and *Hoxa2* genes.

In conclusion, my research has found that *mHotairm1* is a new noncoding transcript in mouse that regulates the expression of *Hoxa1* and *Hoxa2* through the recruitment of MLL1/WDR5 to their chromatin. In NIH 3T3 cells, only sumoylated WDR5 is present in the nuclei and interacts with MLL1 and *mHotairm1*. Following treatment with ATRA, there is an increased sumoylated WDR5 in the nucleus, indicating this modified WDR5 plays a key role in ATRA induced expression of *Hoxa1* and *Hoxa2* (Fig. 6.2)

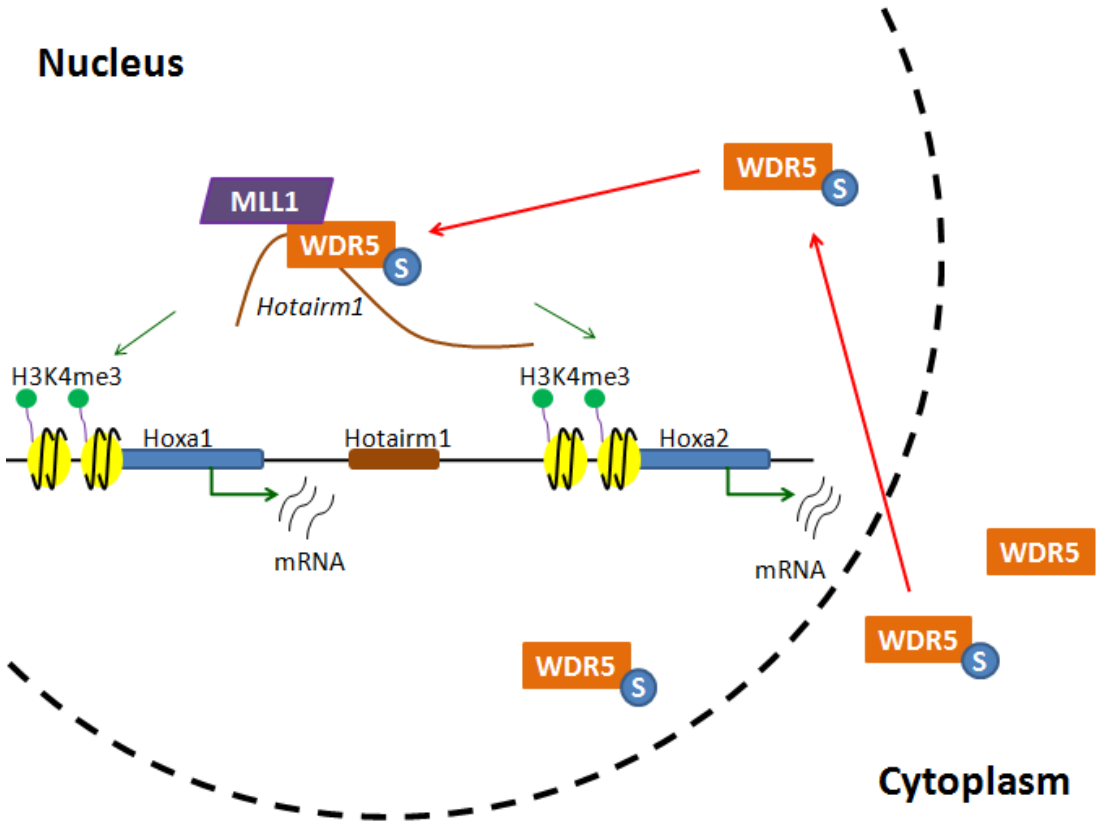


Figure 6.2. Schematic model of *mHotairm1* regulation of *Hoxa1* and *Hoxa2* expression via MLL1/sumoylated WDR5 mediated histone-modification. Sumoylated WDR5 translocates to cell nuclei and binds to the MLL1 histone methytransferase complex and lncRNA *Hotairm1* to induce H3K4me3 occupancy on both *Hoxa1* and *Hoxa2* gene promoters and enhance their expression.

6.4 Future directions

In my experiments, a direct regulation of *Hoxa2* gene expression by the microRNA miR-669b was demonstrated in NIH 3T3 cells. I have also demonstrated expression of miR-669b in the developing mouse palate using real-time PCR which has a complementary expression profile to *Hoxa2* in the developing palate. However, it is not known whether miR-669b has direct role in regulating *Hoxa2* expression in the developing palate. To study the function of miR-669b in mouse palate development, first an overexpression of miR-669b in mouse embryonic palate mesenchyme (MEPM) cell cultures generated from developing mouse palate would be needed. It

would be necessary to show whether *Hoxa2* expression, at both transcriptional and translational level, is downregulated in the MEPM cell cultures after overexpression of miR-669b.

To gain an overview of the spatial distribution profile of miR-669b and *Hoxa2* in all four stages of developing mouse palate, an *in situ* hybridization histochemistry system would need to be undertaken. Locked nucleic acids miRNA probes would be used to achieve higher specificity for the detection of miRNA (Nielsen, 2012; McEwen *et al*., 2016). Mouse palatal organ culture has previously been developed in our lab (Smith *et al*., 2009) and I would use these to determine the role of miR-669b in mouse palatal development by overexpressing miR-669b in mouse palate organ cultures. It is already known that loss of *Hoxa2* function in mice decreases palatal fusion rates and increases in cell proliferation rates (Smith *et al*., 2009). Whether overexpression of miR-669b influencing palatal growth and fusion would also need to be tested.

Further investigation into the function of *mHotairm1* is also required. To study the function of *mHotairm1* in mouse palate development, an improved *in situ* hybridization histochemistry system with higher sensitivity is required so that the distribution of *mHotairm1* in all four stages of developing mouse palate can be examined. It has been suggested that the sensitivity of traditional RNA *in situ* hybridization is too low to visualize lncRNAs in many cases (Raj *et al*., 2008). Recent years, approaches have been made to achieve single molecule RNA *in situ* hybridization, methods including the usage of multiple probes against one target RNA (Raj *et al*., 2008; Bayer *et al*., 2015), usage of chemically modified nucleic acids in probes to strengthen binding specificity (Li *et al*., 2013) and the amplification of fluorescent signals (Larsson *et al*., 2010; Shah *et al*., 2016). These technologies may help better visualize *mHotairm1* expression in developing palate.

Downregulation of *mHotairm1* would be achieved in mouse palatal fusion cultures using siRNAs and the impact of *mHotairm1* on palate fusion and proliferation rates can then be examined. Also *mHotairm1* knockout mice can be generated and the effect of *mHotairm1* on mouse embryo development can be studied in detail. The consequence of loss-of-function of *mHotairm1* on *Hoxa1* and *Hoxa2* gene expression as well as their downstream targets can be examined. The function of *mHotairm1* can also be studied at the cellular level. LncRNAs transcribed from the Hox loci have been reported to be able to regulate cell cycle, cell differentiation and apoptosis (Liu *et al*., 2016; Lian *et al*., 2016; Bertani *et al*., 2011; Zhang *et al*., 2009). Downregulation of *mHotairm1* can be achieved in different cell lines using siRNAs and the influence on cellular activities can be examined.

In my study, the sumoylation of WDR5 was demonstrated by immunoprecipitation but the amino acid where sumoylation may occur has not been identified. Mass spectrometry would be used to determine the amino acid residue that is sumoylated and the key step in this method would be to enrich modified WDR5 protein samples. I previously used immunoprecipitation to accumulate modified WDR5 but unfortunately it was not of sufficient purity with contamination from the IgG heavy chain which also showed up at 50KDa position. In a future experimental design, GST fused MLL1 (amino acid 3810-3963) could be used to accumulate the modified WDR5 protein sample for mass spectrometry as my finding showed there is an interaction between MLL1 and modified WDR5. After identifying the site of sumoylation on WDR5, mutation can be introduced to the amino acid and the role of sumoylation can be examined more thoroughly. The mutated WDR5 can be tagged with GFP and transfected into NIH 3T3 cell lines or MEPM cells to examine the consequence of loss of sumoylation on WDR5, which may affect

the cellular distribution of WDR5, its interaction with lncRNA and MLL1, and the regulation of gene expression.

CHAPTER 7

7. References

Abdelfatah E, Kerner Z, Nanda N and Ahuja N. (2016). Epigenetic therapy in gastrointestinal cancer: the right combination. Therap Adv Gastroenterol. 9:560-579.

Abel T and Zukin RS. (2008). Epigenetic targets of HDAC inhibition in neurodegenerative and psychiatric disorders. Curr Opin in Pharmacol. 8:57–64.

Akin ZN and Nazarali AJ. (2005). Hox Genes and Their Candidate Downstream Targets in the Developing Central Nervous System. Cell Mol Neurobiol. 25:697-741.

Alasti F, Sadeghi A, Sanati MH, Farhadi M, Stollar E, Somers T and Van Camp G. (2008). A mutation in HOXA2 is responsible for autosomal-recessive microtia in an Iranian family. Am J Hum Genet. 82:982-991.

Alharbi RA, Pettengell R, Pandha HS and Morgan R. (2013). The role of HOX genes in normal hematopoiesis and acute leukemia. Leukemia. 27:1000-1008.

Ali A, Veeranki SN and Tyagi S. (2014). A SET-domain-independent role of WRAD complex in cell-cycle regulatory function of mixed lineage leukemia. Nucleic Acids Res. 42:7611-7624.

Ambros V. (2004). The functions of animal microRNAs. Nature. 431:350–355.

Amir RE, Van Den Veyver IB, Wan M, Tran CQ, Francke U and Zoghbi HY. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat. Genet. 23:185–188.

Anderson DM, George R, Noyes MB, Rowton M, Liu W, Jiang R, Wolfe SA, Wilson-Rawls J and Rawls A. (2012). Characterization of the DNA-binding properties of the Mohawk homeobox transcription factor. J Biol Chem. 287:35351-35359.

Ang YS, Tsai SY, Lee DF, Monk J, Su J, Ratnakumar K, Ding J, Ge Y, Darr H, Chang B, Wang J, Rendl M, Bernstein E, Schaniel C and Lemischka IR. (2011). Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. Cell. 145:183-197.

Arand J, Spieler D, Karius T, Branco MR, Meilinger D, Meissner A, Jenuwein T,Xu G, Leonhardt H, Wolf V and Walter J. (2012). In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. PLoS Genet. 8:e1002750.

Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyde B, Gaasterland T, Meyer J and Tuschl T. (2003). The small RNA profile during Drosophila melanogaster development. Dev Cell. 5:337–350.

Archer K, Broskova Z, Bayoumi AS, Teoh JP, Davila A, Tang Y, Su H and Kim IM. (2015). Long Non-coding RNAs as Master Regulators in Cardiovascular Diseases. Int J Mol Sci.16:23651-23667.

Ardekani AM and Naeini MM. (2010). The Role of MicroRNAs in Human Diseases. Avicenna J Med Biotechnol. 2:161-179.

Arlotta P and Hobert O. (2015). Homeotic Transformations of Neuronal Cell Identities. Trends Neurosci. 38:751-762.

Avdic V, Zhang P, Lanouette S, Groulx A, Tremblay V, Brunzelle J and Couture JF. (2011). Structural and biochemical insights into MLL1 core complex assembly. Structure. 19:101–108.

Avvakumov GV, Walker JR, Xue S, Li Y, Duan S, Bronner C, Arrowsmith CH and Dhe-Paganon S. (2008). Structural basis for recognition of hemi- methylated DNA by the SRA domain of human UHRF1. Nature. 455:822-825.

Badeaux AI and Shi Y. (2013). Emerging roles for chromatin as a signal integration and storage platform. Nat Rev Mol Cell Biol. 14:211-224.

Bailey WJ, Kim J, Wagner GP and Ruddle FH. (1997). Phylogenetic reconstruction of vertebrate Hox cluster duplications. Mol Biol Evol. 14:843-853.

Baltz AG, Munschauer M, Schwanhäusser B, Vasile A, Murakawa Y, Schueler M, Yongs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C and Landthaler M. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell. 46: 674–690.

Bandres E, Cubedo E, Agirre X, Malumbres R, Zarate R, Ramirez N, Abajo A, Navarro A, Moreno I, Monzo M and Garcia-Foncillas J. (2006). Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. Mol Cancer. 5:29.

Bandres E, Agirre X, Bitarte N, Ramirez N, Zarate R, Roman-Gomez J, Prosper F and Garcia-Foncillas J. (2009). Epigenetic regulation of microRNA expression in colorectal cancer. Int. J. Cancer. 125: 2737–2743.

Banerjee-Basu S and Baxevanis AD. (2001). Molecular evolution of the homeodomain family of transcription factors. Nucleic Acids Res. 29:3258-3269.

Bannister AJ and Kouzarides T. (1996). The CBP co-activator is a histone acetyltransferase. Nature. 384: 641-643.

Bannister AJ and Kouzarides T. (2011). Regulation of chromatin by histone modifications. Cell Res. 21:381-395.

Bao N, Lye KW and Barton MK. (2004). MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. Dev Cell. 7:653–662.

Barrett LW, Fletcher S and Wilton SD. (2012). Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cell Mol Life Sci. 69:3613-3634.

Bartel DP. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 116: 281-297.

Bartolomei MS and Ferguson-Smith AC. (2011). Mammalian genomic imprinting. Cold Spring Harb Perspect Biol. 3. pii: a002592.

Basu A, Tomar A, Dasari V, Mishra RK and Khosla S. (2016). DNMT3L enables accumulation and inheritance of epimutations in transgenic Drosophila. Sci Rep. 6:19572.

Bayer LV, Batish M, Formel SK and Bratu DP. (2015). Single-Molecule RNA In Situ Hybridization (smFISH) and Immunofluorescence (IF) in the Drosophila Egg Chamber. Methods Mol Biol. 1328:125-136.

Beck S, Faradji F, Brock H and Peronnet F. (2010). Maintenance of Hox gene expression patterns. Adv Exp Med Biol. 689:41-62.

Beh CY, El-Sharnouby S, Chatzipli A, Russell S, Choo SW and White R. (2016). Roles of cofactors and chromatin accessibility in Hox protein target specificity. Epigenetics Chromatin. 9:1.

Bender W. (2008). MicroRNAs in the Drosophila bithorax complex. Genes Dev. 22: 14-19.

Berkyurek AC, Suetake I, Arita K, Takeshita K, Nakagawa A, Shirakawa M and Tajima S. (2014). The DNA methyltransferase Dnmt1 directly interacts with the SET and RING finger-associated (SRA) domain of the multifunctional protein Uhrf1 to facilitate accession of the catalytic center to hemi-methylated DNA. J Biol Chem. 289:379-386.

Berndt A, Wilkinson KA and Henley JM. (2012). Regulation of Neuronal Protein Trafficking and Translocation by SUMOylation. Biomolecules. 2:256-268.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ 3rd, Gingeras TR, Schreiber SL and Lander ES. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell. 120: 169-181.

Bernstein E and Allis CD. (2005). RNA meets chromatin. Genes Dev. 19:1635-1655.

Bertrand N, Roux M, Ryckebusch L, Niederreither K, Dolle P, Moon A, Capecchi M and Zaffran S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. Dev. Biol. 353:266-274.

Bestor TH. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. EMBO J. 11: 2611-2617.

Bertani S, Sauer S, Bolotin E and Sauer F. (2011). The noncoding RNA Mistral activates Hoxa6

and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol Cell. 43:1040-1046.

Bhasin M, Zhang H, Reinherz EL and Reche PA. (2005). Prediction of methylated CpGs in DNA sequences using a support vector machine. FEBS Lett. 579:4302-4308.

Bhatlekar S, Fields JZ and Boman BM. (2014). HOX genes and their role in the development of human cancers. J Mol Med. 92:811–823.

Bird A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. 16:6-21.

Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F and Slany RK. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res. 30, 958–965.

Bhan A, Hussain I, Ansari KI, Kasiri S, Bashyal A and Mandal SS. (2013). Antisense transcript long noncoding RNA (lncRNA) HOTAIR is transcriptionally induced by estradiol. J Mol Biol. 425:3707-22.

Bogdanović O and Veenstra GJ. (2009). DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma. 118:549-565.

Bonasio R and Shiekhattar R. (2014). Regulation of Transcription by Long Noncoding RNAs. Annu Rev Genet. 48:433-455.

Bostick M, Kim JK, Esteve PO, Clark A. Pradhan S and Jacobsen SE. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science. 317: 1760-1764

Bosley TM, Salih MA, Alorainy IA, Oystreck DT, Nester M, Abu-Amero KK, Tischfield MA and Engle EC. (2007). Clinical characterization of the HOXA1 syndrome BSAS variant. Neurology 69, 1245–1253.

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA and Jaenisch R. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature. 441:349-53.

Brehove M, Wang T, North J, Luo Y, Dreher SJ, Shimko JC, Ottesen JJ, Luger K and Poirier MG. (2015). Histone Core Phosphorylation Regulates DNA Accessibility. J Biol Chem. 290:22612-22621.

Breitinger C, Maethner E, Garcia-Cuellar MP and Slany RK. (2012). The homeodomain region controls the phenotype of HOX-induced murine leukemia. Blood. 120:4018-4027.

Brennecke J, Stark A, Russell RB and Cohen SM. (2005). Principles of microRNA-target recognition. PLoS Biol. 3:e85.

Brett D, Hanke J, Lehmann G, Haase S, Delbruck S, Krueger S, Reich J and Bork P. (2000). EST

comparison indicates 38% of human mRNAs contain possible alternative splice forms. FEBS Lett. 474:83-86.

Bridges CB and Morgan TH (1923). The third-chromosome group of mutant characters of Drosophila melanogaster, Vol. 327, Carnegie Institution of Washington Publication, The Lord Baltimore Press, Baltimore, MD, pp.1–251.

Briggs JA, Wolvetang EJ, Mattick JS, Rinn JL and Barry G. (2015). Mechanisms of Long Non-coding RNAs in Mammalian Nervous System Development, Plasticity, Disease, and Evoution. Neuron. 88:861-877.

Brock HW, Hodgson JW, Petruk S and Mazo A. (2009). Regulatory noncoding RNAs at Hox loci. Biochem. Cell Biol. 87:27-34.

Broderick JA and Zamore PD. (2011). MicroRNA therapeutics. Gene Ther. 18:1104-1110.

Brown KK, Viana LM, Helwig CC, Artunduaga MA, Quintanilla-Dieck L, Jarrin P, Osorno G, McDonough B, DePalma SR, Eavey RD, Seidman JG, Seidman CE. (2013). *HOXA2* Haploinsufficiency in Dominant Bilateral Microtia and Hearing Loss. Hum Mutat. 34:1347-1351.

Bürglin TR and Affolter M. (2016). Homeodomain proteins: an update. Chromosoma. 125:497-521.

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A and Rinn JL. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 25:1915–1927.

Canovas S and Ross PJ. (2016). Epigenetics in preimplantation mammalian development. Theriogenology. 86:69-79.

Cao J. (2014). The functional role of long non-coding RNAs and epigenetics. Biol Proced Online. 16:11.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS and Zhang Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science. 298:1039-1043.

Cao R, Tsukada Y and Zhang Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell. 20: 845-854.

Carmell MA, Xuan Z, Zhang MQ and Hannon GJ. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. 16:2733–2742.

Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L, Santoro C, Forrest AR, Carninci P, Biffo S, Stupka E and Gustincich S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature.

491:454-457.

Cheedipudi S, Genolet O and Dobreva G. (2014). Epigenetic inheritance of cell fates during embryonic development. Front Genet. 5:19.

Chen CC, Wang KY and Shen CK. (2013). DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. J Bio Chem. 288:9084-9091.

Chen D, Chen Z, Jin Y, Dragas D, Zhang L, Adjei BS, Wang A, Dai Y and Zhou X. (2013). MicroRNA-99 family members suppress Homeobox A1 expression in epithelial cells. PLoS One. 8:e80625.

Chen D, Yang Y, Cheng X, Fang F, Xu G, Yuan Z, Xia J, Kong H, Xie W, Wang H, Fang M, Gao Y and Xu Y. (2015). Megakaryocytic leukemia 1 directs a histone H3 lysine 4 methyltransferase complex to regulate hypoxic pulmonary hypertension. Hypertension. 65:821-833.

Chen Y and Gorski DH. (2008). Regulation of angiogenesis through a microRNA (miR-130a) that down-regulates antiangiogenic homeobox genes GAX and HOXA5. Blood. 111:1217-1226.

Cheng Y, Jutooru I, Chadalapaka G, Corton JC and Safe S. (2015). The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration. Oncotarget. 6:10840-10852.

Chi Y and Zhou D. (2016). MicroRNAs in colorectal carcinoma - from pathogenesis to therapy. J Exp Clin Cancer Res. 35:43.

Chisaka O, Musci TS and Capecchi MR. (1992). Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene Hox-1.6. Nature 355:516–520.

Chopra VS and Mishra RK. (2006). “Mir” acles in hox gene regulation. BioEssays 28:445–448.

Christensen K, Juel K, Herskind AM and Murray JC. (2004). Long Term Follow Up Study of Survival Associated with Cleft Lip and Palate at Birth. B. M. J. 328: 1405-1408.

Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR and Goodman RH. (1993). Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature. 365: 855-859.

Colamaio M, Puca F, Ragozzino E, Gemei M, Decaussin-Petrucci M, Aiello C, Bastos AU, Federico A, Chiappetta G, Del Vecchio L, Torregrossa L, Battista S and Fusco A. (2015). miR-142-3p down-regulation contributes to thyroid follicular tumorigenesis by targeting ASH1L and MLL1. J Clin Endocrinol Metab. 100:E59-69.

Cordero DR, Brugmann S, Chu Y, Bajpai R, Jame M and Helms JA. (2011). Cranial neural crest cells on the move: their roles in craniofacial development. Am J Med Genet A. 155A:270-279.

Cosgrove MS and Patel A. (2010). Mixed lineage leukemia: a structure-function perspective of

the MLL1 protein. FEBS J. 277:1832-1842.

Couture JF, Collazo E and Trievel RC. (2006). Molecular recognition of histone H3 by the WD40 protein WDR5. Nat Struct Mol Biol. 13:698-703.

Cox TC, Camci ED, Vora S, Luquetti DV and Turner EE. (2014). The genetics of auricular development and malformation: new findings in model systems driving future directions for microtia research. Eur J Med Genet. 57:394-401.

Cunningham TJ and Duester G. (2015). Mechanisms of retinoic acid signalling and its roles in organ and limb development. Nat Rev Mol Cell Biol. 16:110-123.

Daftary GS and Taylor HS. (2006). Endocrine regulation of HOX genes. Endocr Rev. 27:331-355.

Dai X, Guo W, Zhan C, Liu X, Bai Z and Yang Y. (2015). WDR5 Expression Is Prognostic of Breast Cancer Outcome. PLoS One. 10:e0124964.

Darda L, Hakami F, Morgan R, Murdoch C, Lambert DW and Hunter KD. (2015). The role of HOXB9 and miR-196a in head and neck squamous cell carcinoma. PLoS One. 10:e0122285.

Das Gupta K, Shakespear MR, Iyer A, Fairlie DP and Sweet MJ. (2016). Histone deacetylases in monocyte/macrophage development, activation and metabolism: refining HDAC targets for inflammatory and infectious diseases. Clin Transl Immunology. 5:e62.

De Kumar B and Krumlauf R. (2016). *HOX*s and lincRNAs: Two sides of the same coin. Sci Adv. 2:e1501402.

Deaton AM and Bird A. (2011). CpG islands and the regulation of transcription. Genes Dev. 25:1010-1022.

Deiuliis JA. (2016). MicroRNAs as regulators of metabolic disease: Pathophysiologic significance and emerging role as biomarkers and therapeutics. Int J Obes (Lond). 40:88-101.

Deprez PM, Nichane MG, Rousseaux P, Devogelaer JP, Chappard D, Lengele´ BG, Rezsöhazy R and Nyssen-Behets C. (2012). Postnatal growth defect in mice upon persistent Hoxa2 expression in the chondrogenic cell lineage. Differentiation. 83:158-167.

Deprez PM, Nichane MG, Lengele´ BG, Rezsöhazy R and Nyssen-Behets C. (2013). Molecular study of a Hoxa2 gain-of-function in chondrogenesis: a model of idiopathic proportionate short stature. Int J Mol Sci. 14:20386-20398.

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J and Guigó R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 22:1775-1789.

Di Giammartino DC, Nishida K and Manley JL. (2011). Mechanisms and consequences of alternative polyadenylation. Mol Cell. 43:853-866.

Dias J, Van Nguyen N, Georgiev P, Gaub A, Brettschneider J, Cusack S, Kadlec J and Akhtar A. (2014). Structural analysis of the KANSL1/WDR5/KANSL2 complex reveals that WDR5 is required for efficient assembly and chromatin targeting of the NSL complex. Genes Dev. 28:929-942.

Doench JG and Sharp PA. (2004). Specificity of microRNA target selection in translational repression. Genes Dev. 18:504–511.

Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L and Schübeler D. (2015). Competition between DNA methylation and transcription factors determines binding of NRF1. Nature. 528:575-579.

Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD and Roeder RG. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat StructMol Biol. 13:713-719.

Dovey OM, Foster CT and Cowley SM. (2010). Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. Proc Natl Acad Sci U S A. 107:8242–8247.

Du Q, Luu PL, Stirzaker C and Clark SJ. (2015). Methyl-CpG-binding domain proteins: readers of the epigenome. Epigenomics. 7:1051-1073.

Duboule D. (2007). The rise and fall of Hox gene clusters. Development. 134:2549-2560.

Dugas DV and Bartel B. (2004). MicroRNA regulation of gene expression in plants. Curr Opin Plant Biol. 7:512–520.

Eckner R, Ewen ME, Newsome D, Gerdes M, DeCaprio JA, Lawrence JB and Livingston DM. (1994). Molecular cloning and functional analysis of the adenovirus E1A-associated 300-KD protein (p300) reveals a protein with properties of a transcriptional adaptor. Genes Dev. 8: 869-884.

Eichhorn SW, Guo H, McGeary SE, Rodriguez-Mias RA, Shin C, Baek D, Hsu SH, Ghoshal K, Villén J and Bartel DP. (2014). mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. Mol. Cell. 56:104-115.

Elia L and Condorelli G. (2015). RNA (Epi)genetics in cardiovascular diseases. J Mol Cell Cardiol. 89:11-16.

Ernst P and Vakoc CR. (2012). WRAD: enabler of the SET1-family of H3K4 methyltransferases. Brief Funct Genomics. 11:217-226.

Eun Jeoung L, Sung Hee H, Jaesun C, Sung Hwa S, Kwang Hum Y, Min Kyoung K, Tae Yoon P and Sang Sun K. (2008). Regulation of glycogen synthase kinase 3beta functions by

modification of the small ubiquitin-like modifier. Open Biochem J. 2:67-76.

Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C, Volinia S, Guler G, Morrison CD, Chan KK, Marcucci G, Calin GA, Huebner K and Croce CM. (2007). MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proc Natl Acad Sci USA. 104:15805–15810.

Fan G and Hutnick L. (2005). Methyl-CpG binding proteins in the nervous system. Cell Res. 15:255–261.

Fan J, Krautkramer KA, Feldman JL and Denu JM. (2015). Metabolic regulation of histone post-translational modifications. ACS Chem Biol. 10:95-108.

Fang K, Liu P, Dong S, Guo Y, Cui X, Zhu X, Li X, Jiang L, Liu T and Wu Y. (2016). Magnetofection based on superparamagnetic iron oxide nanoparticle-mediated low lncRNA HOTAIR expression decreases the proliferation and invasion of glioma stem cells. Int J Oncol. 49:509-518.

Fang X, Poulsen RR, Wang-Hu J, Shi O, Calvo NS, Simmons CS, Rivkees SA and Wendler CC. (2016). Knockdown of DNA methyltransferase 3a alters gene expression and inhibits function of embryonic cardiomyocytes. FASEB J. pii: fj.201600346R.

Faust C, Schumacher A, Holdener B and Magnuson T. (1995). The eed mutation disrupts anterior mesoderm production in mice. Development. 121: 273–285.

Favier B and Dollé P. (1997). Developmental functions of mammalian Hox genes. Mol Hum Reprod. 3:115-131.

Felekkis K, Touvana E, Stefanou Ch and Deltas C. (2010). microRNAs: a newly described class of encoded molecules that play a role in health and disease. Hippokratia. 14:236-240.

Feng J, Fouse S, and Fan G. (2007). Epigenetic Regulation of Neural Gene Expression and Neuronal Function. Pediatr Res. 61 (5 Pt 2): 58R-63R.

Ferguson MW. (1988). Palate Development. Development. 103, 41-60.

Fries GR, Li Q, McAlpin B, Rein T, Walss-Bass C, Soares JC and Quevedo J. (2016). The role of DNA methylation in the pathophysiology and treatment of bipolar disorder. Neurosci Biobehav Rev. 68:474-488.

Froberg JE, Yang L and Lee JT. (2013). Guided by RNAs: X-inactivation as a model for lncRNA function. J Mol Biol. 425:3698-3706.

Furlan G and Rougeulle C. (2016). Function and evolution of the long noncoding RNA circuitry orchestrating X-chromosome inactivation in mammals. Willey Interdiscip Rev RNA. 7:702-722.

Garaulet DL, Castellanos MC, Bejarano F, Sanfilippo P, Tyler DM, Allan DW, Sánchez-Herrero

E and Lai EC. (2014). Homeotic function of Drosophila Bithorax-Complex miRNAs mediates fertility by restricting multiple Hox Genes and TALE cofactors in the CNS. Dev Cell. 29:635-648.

Garaulet DL and Lai EC. (2015). Hox miRNA regulation within the Drosophila Bithorax complex: Patterning behavior. Mech Dev. 138 Pt 2:151-159.

Gaunt SJ, Sharpe PT and Duboule D. (1988). Spatially restricted domains of homeo-gene transcripts in mouse embryos: relation to a segmented body plan. Development. 104: 169-179.

Gavalas A, Davenne M, Lumsden A, Chambon P and Rijli FM. (1997). Role of Hoxa-2 in axon pathfinding and rostral hindbrain patterning. Development. 124:3693-3702.

Ge Y, Yan X, Jin Y, Yang X, Yu X, Zhou L, Han S, Yuan Q and Yang M. (2015). MiRNA-192 [corrected] and miRNA-204 Directly Suppress lncRNA HOTTIP and Interrupt GLS1-Mediated Glutaminolysis in Hepatocellular Carcinoma. PLoS Genet. 11:e1005726.

Gebauer F and Hentze MW. (2004). Molecular mechanisms of translational control. Nat Rev Mol Cell Biol. 5:827–835.

Gehring WJ. (1993). Exploring the homeobox. Gene. 135:215–221.

Gehring WJ, Qian YG, Biiieter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D; Affolter M; Otting G and Wuhrich K. (1994). Homeodomain-DNA Recognition. Cell. 78:211-223.

Geisler SJ and Paro R. (2015). Trithorax and Polycomb group-dependent regulation: a tale of opposing activities. Development. 142:2876-2887.

Gendron-Maguire M, Mallo M, Zhang M and Gridley T. (1993). Hoxa-2 mutant mice exhibit homeotic transformation of skeletal elements derived from cranial neural crest. Cell. 75:1317-1331.

Giraldez AJ, Cinalli RM, Glasner ME, Enright AJ, Thomson JM, Baskerville S, Hammond SM, Bartel DP and Schier AF. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. Science. 308: 833–838.

Godlewski J, Nowicki MO, Bronisz A, Williams S, Otsuki A, Nuovo G, Raychaudhury A, Newton HB, Chiocca EA and Lawler S. (2008). Targeting of the Bmi-1 oncogene /stem cell renewal factor by microRNA-128 inhibits glioma proliferation and self-renewal. Cancer Res 68:9125–9130.

Goldberg AD, Allis CD and Bernstein E. (2007). Epigenetics: A Landscape Takes Shape. Cell. 128:635-638.

Goodman RH and Smolik S. (2000). CBP/p300 in cell growth, transformation, and development. Genes Dev. 14:1553-1577.

Goodrich L, Panning B and Leung KN. (2016). Activators and repressors: A balancing act for X-inactivation. Semin Cell Dev Biol. pii: S1084-9521(16)30129-X.

Gori F, Divieti P and Demay MB. (2001). Cloning and Characterization of a Novel WD-40 Repeat Protein That Dramatically Accelerates Osteoblastic Differentiation. J Biol Chem. 276:46515-46522.

Gori F, Friedman L and Demay MB (2005). Wdr5, a novel WD repeat protein, regulates osteoblast and chondrocyte differentiation in vivo. J Musculoskelet Neuronal Interact. 5:338-339

Gori F, Friedman LG and Demay MB. (2006). Wdr5, a WD-40 protein, regulates osteoblast differentiation during embryonic bone development. Dev Biol. 295:498-506.

Gori F, Zhu ED and Demay MB. (2009). Perichondrial expression of Wdr5 regulates chondrocyte proliferation and differentiation. Dev. Biol. 329:36-43.

Graves HK, Wang P, Lagarde M, Chen Z and Tyler JK. (2016). Mutations that prevent or mimic persistent post-translational modifications of the histone H3 globular domain cause lethality and growth defects in Drosophila. Epigenetics Chromatin. 9:9.

Greco S, Zaccagnini G, Perfetti A, Fuschi P, Valaperta R, Voellenkle C, Castelvecchio S, Gaetano C, Finato N, Beltrami AP, Menicanti L and Martelli F. (2016). Long noncoding RNA dysregulation in ischemic heart failure. J Transl Med. 14:183.

Green D, Dalmay T and Fraser WD. (2015). Role of miR-140 in embryonic bone development and cancer. Clin Sci (Lond). 129:863-873.

Grimmel M, Backhaus C and Proikas-Cezanne T. (2015). WIPI-Mediated Autophagy and Longevity. Cells. 4:202-217.

Grun D, Wang YL, Langenberger D, Gunsalus KC and Rajewsky N. (2005). microRNA target predictions across seven Drosophila species and comparison to mammalian targets. PLoS Comput Biol. 1:e13.

Guan JS, Haggarty SJ, Giacometti E, Dannenberg JH, Joseph N, Gao J, Nieland TJ, Zhou Y, Wang X, Mazitschek R, Bradner JE, DePinho RA, Jaenisch R, Tsai LH. (2009). HDAC2 negatively regulates memory formation and synaptic plasticity. Nature 459:55–60.

Gummalla M, Maeda RK, Castro Alvarez JJ, Gyurkovics H, Singari S, Edwards KA, Karch F and Bender W. (2012). abd-A regulation by the iab-8 noncoding RNA. PLoS Genet. 8:e1002720.

Ha M and Kim VN. (2014). Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15:509–524.

Han B, Lian L, Li X, Zhao C, Qu L, Liu C, Song J and Yang N. (2016). Chicken gga-miR-130a targets HOXA3 and MDFIC and inhibits Marek's disease lymphoma cell proliferation and migration. Mol Biol Rep. 43:667-676.

Han L, Witmer PD, Casey E, Valle D and Sukumar S. (2007). DNA Methylation Regulates MicroRNA Expression. Cancer Biol Ther. 6:1284-1288.

Han Z, Guo L, Wang H, Shen Y, Deng XW and Chai J. (2006). Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD-40 protein WDR5. Mol Cell. 22:137-144.

He X, Yan YL, DeLaurier A and Postlethwait JH. (2011a). Observation of miRNA gene expression in zebrafish embryos by in situ hybridization to microRNA primary transcripts. Zebrafish. 8:1-8.

He X, Yan YL, Eberhart JK, Herpin A, Wagner TU, Schartl M and Postlethwait JH. (2011b). miR-196 regulates axial patterning and pectoral appendage initiation. Dev. Biol. 357:463-477.

Hellman A and Chess A. (2007). Gene body-specific methylation on the active X chromosome. Science. 315:1141–1143.

Helmbacher F, Pujades C, Desmarquet C, Frain M, Rijli FM, Chambon P and Charnay P. (1998). Hoxa1 and Krox-20 synergize to control the development of rhombomere3. Development. 125, 4739–4748.

Holliday R. (2006) Epigenetics: a historical overview. Epigenetics. 1:76-80.

Holve S, Friedman B, Hoyme HE, Tarby TJ, Johnstone SJ, Erickson RP, Clericuzio CL and Cunniff C. (2003). Athabascan brainstem dysgenesis syndrome. Am. J. Med. Genet. A 120: 169–173.

Hornstein E, Mansfield JH, Yekta S, Hu JK, Harfe BD, McManus MT, Baskerville S, Bartel DP and Tabin CJ. (2005). The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. Nature. 438:671-674.

Hu JL, Zhou BO, Zhang RR, Zhang KL, Zhou JQ and Xu GL. (2009). The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. Proc. Natl. Acad. Sci. USA. 106:22187–22192.

Hu R, Liu W, Li H, Yang L, Chen C, Xia ZY, Guo LJ, Xie H, Zhou HD, Wu XP and Luo XH. (2011). A Runx2/ miR-3960/miR-2861 regulatory feedback loop during mouse osteoblast differentiation. J Biol Chem. 286:12328-12339.

Hueber SD and Lohmann I. (2008). Shaping segments: Hox gene function in the genomic age. BioEssays. 30:965–979.

Huntzinger E and Izaurralde E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat. Rev. Genet. 12: 99-110.

Hyllus D, Stein C, Schnabel K, Schiltz E, Imhof A, Dou Y, Hsieh J and Bauer UM. (2007). PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. Genes Dev. 21:3369-3380.

Ipsaro JJ and Joshua-Tor L. (2015). From guide to target: molecular insights into eukaryotic RNA-interference machinery. Nat Struct Mol Biol. 22:20–28.

Jeltsch A, Ehrenhofer-Murray A, Jurkowski TP, Lyko F, Reuter G, Ankri S, Nellen W, Schaefer M and Helm M. (2016). Mechanism and biological role of Dnmt2 in Nucleic Acid Methylation. RNA Biol. 1-16.

Jenuwein T and Allis CD. (2001). Translating the histone code. Science. 293:1074–1080.

Jiang H, Shukla A, Wang X, Chen WY, Bernstein BE and Roeder RG. (2011). Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. Cell. 144:513-525.

Jin JZ and Ding J. (2006). Analysis of cell migration, transdifferentiation and apoptosis during mouse secondary palate fusion. Development. 133:3341-3347.

Jin K and Sukumar S. (2016). HOX genes: Major actors in resistance to selective endocrine response modifiers. Biochim Biophys Acta. 1865:105-110.

Jonas S and Izaurralde E. (2015). Towards a molecular understanding if microRNA-mediated gene silencing. Nat Rev Genet. 16:421-433.

Jones PA. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 13:484-492.

Jones A and Wang H. (2010). Polycomb repressive complex 2 in embryonic stem cells: an overview. Protein Cell. 1: 1056–1062.

Jørgensen S, Schotta G and Sørensen CS. (2013). Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. Nucleic Acids Res. 41:2797-2806.

Jung HJ and Suh Y. (2015). Regulation of IGF -1 signaling by microRNAs. Front Genet. 5:472.

Jurgens G. (1985). A group of genes controlling spatial expression of the bithorax complex in Drosophila. Nature. 316: 153–155.

Jurkowska RZ, Jurkowski TP and Jeltsch A. (2011) Structure and Function of Mammalian DNA Methyltransferases. ChemBioChem. 12: 206-222.

Kadoch C, Copeland RA and Keilhack H. (2016). PRC2 and SWI/SNF chromatin remodeling complex in health and disease. Biochemistry. 55:1600-1614.

Kanzler B, Kuschert SJ, Liu Y-H and Mallo M. (1998). Hoxa2 restricts the chondrogenic domain and inhibits bone formation during development of the branchial area. Development. 125: 2587–2597.

Kashyap V, Gudas LJ, Brenet F, Funk P, Viale A and Scandura JM. (2011). Epigenetic reorganization of the clustered hox genes in embryonic stem cells induced by retinoic acid. J Biol

Chem. 286:3250-3260.

Kayam G, Kohl A, Magen Z, Peretz Y, Weisinger K, Bar A, Novikov O, Brodski C and Sela-Donenfeld D. (2013). A novel role for Pax6 in the segmental organization of the hindbrain. Development. 140(10):2190-2202.

Ke XS, Liu CM, Liu DP and Liang CC. (2003). MicroRNAs: Key participants in gene regulatory networks. Curr Opin Chem Biol. 7:516-523.

Khraiwesh B, Arif MA, Seumel GI, Ossowski S, Weigel D, Reski R and Frank W. (2009). Transcriptional Control of Gene Expression by MicroRNAs. Cell. 140: 111–122.

Kim HJ, Kim YH, Lee DS, Chung JK and Kim S. (2008). In vivo imaging of functional targeting of miR-221 in papillary thyroid carcinoma. J Nucl Med. 49:1686-1693.

Kim JY, Banerjee T, Vinckevicius A, Luo Q, Parker JB, Baker MR, Radhakrishnan I, Wei JJ, Barish GD and Chakravarti D. (2014). A role for WDR5 in integrating threonine 11 phosphorylation to lysine 4 methylation on histone H3 during androgen signaling and in prostate cancer. Mol Cell. 54:613-625.

Kissinger CF, Martin-Blanco E and Pabo CO. (1990). Crystal Structure of an engrailed Homeodomain-DNA Complex at 2.8 A Resolution: A Framework for Understanding Homeodomain-DNA Interactions. Cell. 63:579-590.

Kitazawa T, Fujisawa K, Narboux-Neme N, Arima Y, Kawamura Y, Inoue T, Wada Y, Kohro T, Aburatani H, Kodama T, Kim K, Sato T, Uchijima Y, Maeda K, Miyagawa-Tomita S, Minoux M, Rijli FM, Levi G, Kurihara Y, Kurihara H. (2013). Distinct effects of Hoxa2 overexpression in cranial neural crest populations reveal that the mammalian hyomandibular-ceratohyal boundary maps within the styloid process. Dev. Biol. 402:162-174.

Klose RJ and Bird AP. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci. 31:89-97.

Kocerha J, Kauppinen S and Wahlestedt C. (2009). microRNAs in CNS Disorders. Neuromol Med. 11:162–172.

Kong L, Zhou X, Wu Y, Wang Y, Chen L, Li P, Liu S, Sun S, Ren Y, Mei M, Wang X and Zhang L. (2015). Targeting HOTAIR Induces Mitochondria Related Apoptosis and Inhibits Tumor Growth in Head and Neck Squamous Cell Carcinoma in vitro and in vivo. Curr Mol Med. 15:952-960.

Krishnan S, Horowitz S and Trievel RC. (2011). Structure and Function of Histone H3 Lysine 9 Methyltransferases and Demethylases. ChemBioChem 12: 254 – 263.

Kumar A and Zhang KY. (2015). Advances in the development of SUMO specific protease (SENP) inhibitors. Comput Struct Biotechnol J. 13:204-211.

Kumar P and Nazarali AJ. (2001). Characterization of Hoxd1 protein-DNA-binding specificity using affinity chromatography and random DNA oligomer selection. Cell. Mol. Neurobiol. 21:369–388.

Kurihara Y, Kawamura Y, Uchijima Y, Amamo T, Kobayashi H, Asano T and Kurihara H. (2008). Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. Dev Biol. 313:335-346.

Laget S, Miotto B, Chin HG, Estève PO, Roberts RJ, Pradhan S and Defossez PA. (2014). MBD4 cooperates with DNMT1 to mediate methyl-DNA repression and protects mammalian cells from oxidative stress. Epigenetics. 9:546–556.

Lagos-Quintana M, Rauhut R, Lendeckel W and Tuschl T. (2001). Identification of novel genes coding for small expressed RNAs. Science. 294: 853–858.

Lalonde ME, Cheng X and Côté J. (2014). Histone target selection within chromatin: an exemplary case of teamwork. Genes Dev. 28:1029-1041.

Lamoliatte F, Caron D, Durette C, Mahrouche L, Maroui MA, Caron-Lizotte O, Bonneil E, Chelbi-Alix MK and Thibault P. (2014). Large-scale analysis of lysine SUMOylation by SUMO remnant immunoaffinity profiling. Nat Commun. 5:5409.

Lan Y, Xu J and Jiang R. (2015). Cellular and Molecular Mechanisms of Palatogenesis. Curr Top Dev Biol. 115:59-84.

Lappin TR, Grier DG, Thompson A and Halliday HL. (2006). HOX genes: seductive science, mysterious mechanisms. Ulster Med J. 75:23-31.

Larsson C, Grundberg I, Söderberg O and Nilsson M. (2010). In situ detection and genotyping of individual mRNA molecules. Nature methods. 7: 395-397.

Lee J. (2000). Disruption of imprinted X inactivation by parent-of-origin effects at Tsix. Cell. 103:17–27.

Lee JT. (2012). Epigenetic Regulation by Long Noncoding RNAs. Science. 338:1435-1439.

Lee M, Kim HJ, Kim SW, Park SA, Chun KH, Cho NH, Song YS and Kim YT. (2016). The long non-coding RNA HOTAIR increases tumour growth and invasion in cervical cancer by targeting the Notch pathway. Oncotarget. doi: 10.18632/oncotarget.10065.

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, *et al*. (2006). Control of developmental regulators by polycomb in human embryonic stem cells. Cell. 125: 301–313.

Lee RC, Feinbaum RL and Ambros V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 75: 843–854.

Lewis EB. (1978). A gene complex controlling segmentation in Drosophila. Nature. 276:565–

570.

Li D and Roberts R (2001). WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. Cell. Mol. Life Sci. 58: 2085–2097

Li D, Feng J, Wu T, Wang Y, Sun Y, Ren J and Liu M. (2013). Long intergenic noncoding RNA HOTAIR is overexpressed and regulates PTEN methylation in laryngeal squamous cell carcinoma. Am J Patrol. 182:64-70.

Li E, Bestor TH and Jaenisch R. (1992). Targeted Mutation of the DNA Methyltransferase Gene Results in Embryonic Lethality. Cell. 69: 915 –926.

Li J, Yang S, Su N, Wang Y, Yu J, Qiu H and He X. (2016). Overexpression of long non-coding RNA HOTAIR leads to chemoresistance by activating the Wnt/β-catenin pathway in human ovarian cancer. 37:2057-2065.

Li X, Li L, Pandey R, Byun JS, Gardner K, Qin Z, Dou Y. (2012). The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. Cell Stem Cell. 11:163-178.

Li Q, Zhu F and Chen P. (2012). miR-7 and miR-218 epigenetically control tumor suppressor genes RASSF1A and Claudin-6 by targeting HoxB3 in breast cancer. Biochem Biophys Res Commun. 424:28-33.

Li W, Notani D, Ma Q, Tanasa B, Nunez E, Chen AY, Merkurjev D, Zhang J, Ohgi K, Song X, Oh S, Kim HS, Glass CK and Rosenfeld MG. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 498:516-520.

Li X, Xiao B and Chen XS. (2016). DNA Methylation: a New Player in Multiple Sclerosis. Mol Neurobiol. doi:10.1007/s12035-016-9966-3.

Li Z, Huang H, Chen P, He M, Li Y, Arnovitz S, Jiang X, He C, Hyjek E, Zhang J, Zhang Z, Elkahloun A, Cao D, Shen C, Wunderlich M, Wang Y, Neilly MB, Jin J, Wei M, Lu J, Valk PJ, Delwel R, Lowenberg B, Le Beau MM, Vardiman J, Mulloy JC, Zeleznik-Le NJ, Liu PP, Zhang J and Chen J. (2012). miR-196b directly targets both HOXA9/MEIS1 oncogenes and FAS tumour suppressor in MLL-rearranged leukaemia. Nat Commun. 3:688.

Li Z, Xiao J, Hu K, Wang G, Li M, Zhang J and Cheng G. (2015a). FBXW7 acts as an independent prognostic marker and inhibits tumor growth in human osteosarcoma. Int J Mol Sci. 16:2294-2306.

Li Z, Zhao L and Wang Q. (2016). Overexpression of long non-coding RNA HOTTIP increases chemoresistance of osteosarcoma cell by activating the Wnt/β-catenin pathway. Am J Transl Res. 8:2385-2393.

Li Z, Zhao X, Zhou Y, Liu Y, Zhou Q, Ye H, Wang Y, Zeng J, Song Y, Gao W, Zheng S, Zhuang B, Chen H, Li W, Li H, Li H, Fu Z and Chen R. (2015b). The long non-coding RNA HOTTIP promotes progression and gemcitabine resistance by regulating HOXA13 in

pancreatic cancer. J Transl Med.13:84.

Lian Y, Ding J, Zhang Z, Shi Y, Zhu Y, Li J, Peng P, Wang J, Fan Y, De W and Wang K. (2016). The long noncoding RNA HOXA transcript at the distal tip promotes colorectal cancer growth partially via silencing of p21 expression. Tumour Biol. 37:7431-7440.

Liu J, Wang L, Su Z, Wu W, Cai X, Li D, Hou J, Pei D and Pan G. (2014). A reciprocal antagonism between miR-376c and TGF-β signaling regulates neural differentiation of human pluripotent stem cells. FASEB J. 28:4642-4656.

Liu Y, Wang B, Liu X, Lu L, Luo F, Lu X, Shi L, Xu W and Liu Q. (2016a). Epigenetic silencing of p21 by long non-coding RNA HOTAIR is involved in the cell cycle disorder induced by cigarette smoke extract. Toxicol Lett. 240:60-67.

Liu Y, Zhang B, Kuang H, Korakavi G, Lu LY and Yu X. (2016b). Zinc Finger Protein 618 Regulates the Function of UHRF2 (Ubiquitin-like with PHD and Ring Finger Domains 2) as a Specific 5-Hydroxymethylcytosine Reader. J Biol Chem. 291:13679-13688.

Liu Y, Zheng W, Song Y, Ma W and Yin H. (2013). Low expression of miR-196b enhances the expression of BCR-ABL1 and HOXA9 oncogenes in chronic myeloid leukemogenesis. PLoS One. 8:e68442.

Lomvardas S and Maniatis T. (2016). Histone and DNA Modifications as Regulators of Neuronal Development and Function. Cold Spring Harb Perspect Biol. 8:pii:a024208.

Luo T, Cui S, Bian C, Yu X. (2013). Uhrf2 is important for DNA damage response in vascular smooth muscle cells. Biochem Biophys Res Commun. 441:65-70.

Lv L, Li Y, Deng H, Zhang C, Pu Y, Qian L, Xiao J, Zhao W, Liu Q, Zhang D, Wang Y, Zhang H, He Y and Zhu J. (2015). MiR-193a-3p promotes the multi-chemoresistance of bladder cancer by targeting the HOXC9 gene. Cancer Lett. 357:105-113.

Lv W, Guo X, Wang G, Xu Y and Kang J. (2014). Histone deacetylase 1 and 3 regulate the mesodermal lineage commitment of mouse embryonic stem cells. PLoS One. 9:e113262.

Lyon M. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature. 190:372–373.

Ma L, Teruya-Feldstein J and Weinberg RA. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature. 449: 682–688.

Macfarlane LA and Murphy PR. (2010). MicroRNA: Biogenesis, Function and Role in Cancer. Curr Genomics. 11:537-561.

Maes OC, Chertkow HM, Wang E and Schipper HM. (2009). MicroRNA: Implications for Alzheimer Disease and other Human CNS Disorders. Current Genomics. 10: 154-168.

Mahajan R, Delphin C, Guan T, Gerace L and Melchior F. (1997). A small ubiquitin-related

polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell. 88:97-107.

Makki N and Capecchi MR. (2010). *Hoxa1* lineage tracing indicates a direct role for *Hoxa1* in the development of the inner ear, the heart, and the third rhombomere. Dev Biol. 15;341(2):499-509.

Makki N and Capecchi MR. (2011). Identification of novel Hoxa1 downstream targets regulating hindbrain, neural crest and inner ear development. Dev Biol. 357:295-304.

Makki N and Capecchi MR. (2012). Cardiovascular defects in a mouse model of *HOXA1* syndrome. Hum Mol Genet. 21: 26–31.

Mallo M and Alonso CR. (2013). The regulation of Hox gene expression during animal development. Development. 140:3951-3963.

Mann RS and Affolter M. (1998). Hox proteins meet more partners. Curr. Opin.Genet. Dev. 8:423–429.

Mann RS, Lelli KM and Joshi R. (2009). Hox specificity: unique roles for cofactors and collaborators. Curr Top Dev Biol. 88:63-101.

Martin CH, Mayeda CA, Davis CA, Ericsson CL, Knafels JD, Mathog DR, Celniker SE, Lewis EB and Palazzolo MJ. (1995). Complete sequence of the bithorax complex of Drosophila. Proc. Natl. Acad. Sci. 92: 8398–8402.

Martinez-Ceballos E and Gudas LJ. (2008). Hoxa1 is required for the retinoic acid–induced differentiation of embryonic stem cells into neurons. J Neurosci Res. 86:2809-2819.

Massip L, Ectors F, Deprez P, Maleki M, Behets C, Lengele´ B, Delahaut P, Picard J, Rezsohazy R. (2007). Expression of Hoxa2 in cells entering chondrogenesis impairs overall cartilage development. Differentiation. 75:256-267.

Matunis MJ, Coutavas E and Blobel G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. J Cell Biol. 135:1457-1470.

McEwen TJ, Yao Q, Yun S, Lee CY and Bennett KL. (2016). Small RNA in situ hybridization in Caenorhabditis elegans, combined with RNA-seq, identifies germline-enriched microRNAs. Dev Biol. 418:248-257.

McGinnis W and Krumlauf R. (1992). Homeobox genes and axial patterning. Cell. 68:283–302.

McHugh CA, Chen CK, Chow A, Surka CF, Tran S, McDorel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A, Sweredoski MJ, Shishkin AA, Su J, Lander ES, Hess S, Plath K and Guttman M. (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. 521:232–236.

Mercer TR and Mattick JS. (2013). Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol. 20:300-307.

Miguez A, Ducret S, Meglio TD, Parras C, Hmidan H, Haton C, Sekizar S, Mannioui A, Vidal M, Kerever A, Nyabi O, Haigh J, Zalc B, Rijli FM, Thomas JL. (2012). Opposing Roles for *Hoxa2* and *Hoxb2* in Hindbrain Oligodendrocyte Patterning. J Neurosci. 32:17172-17185.

Minoux M, Kratochwil CF, Ducret S, Amin S, Kitazawa T, Kurihara H, Bobola N, Vilain N and Rijli FM. (2013). Mouse *Hoxa2* mutations provide a model for microtia and auricle duplication. Development. 140:4386-4397.

Minoux M and Rijli FM. (2010). Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. Development. 137:2605-2621.

Miura P, Shenker S, Andreu-Agullo C, Westholm JO and Lai EC. (2013). Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. Genome Res. 23:812–825.

Monks DC, Jahangir A, Shanske AL, Samanich J, Morrow BE and Babcock M. (2010). Mutational analysis of HOXA2 and SIX2 in a Bronx population with isolated microtia. Int J Pediatr Otorhinolaryngol. 74:878-882.

Montavon T and Duboule D. (2013). Chromatin organization and global regulation of Hox gene clusters. Philos Trans R Soc Lond B Biol Sci. 368:20120367.

Montgomery RL, Potthoff MJ, Haberland M, Qi X, Matsuzaki S, Humphries KM, Richardson JA, Bassel-Duby R and Olson EN. (2008). Maintenance of cardiac energy metabolism by histone deacetylase 3 in mice. J Clin Invest. 118:3588–3597.

Moore LD, Le T and Fan G. (2013). DNA methylation and its basic function. Neuropsychopharmacology. 38:23-38.

Morey L, Santanach A and Di Croce L. (2015). Pluripotency and Epigenetic Factors in Mouse Embryonic Stem Cell Fate Regulation. Mol Cell Biol. 35:2716-2728.

Mu X, Yan S, Fu C, Wei A. (2015). The Histone Acetyltransferase MOF Promotes Induces Generation of Pluripotent Stem Cells. Cell Reprogram. 17: 259-267.

Mujahid S, Nielsen HC and Volpe MV. (2013). MiR-221 and miR-130a regulate lung airway and vascular development. PLoS One. 8:e55911.

Murphy P and Hill RE. (1991). Expression of the mouse labial-like homeobox-containing genes, Hox 2.9 and Hox 1.6, during segmentation of the hindbrain. Development. 111: 61–74.

Nagy PL, Griesenbeck J, Kornberg RD and Cleary ML. (2002). A trithoraxgroup complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. Proc Natl Acad Sci USA. 99:90–94.

Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A,

Croce CM and Canaani E. (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell. 10:1119-1128.

Naguibneva I, Ameyar-Zazoua M, Polesskaya A, Ait-Si-Ali S, Groisman R, Souidi M, Cuvellier S and Harel-Bellan A. (2006). The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. Nature Cell Biology. 8:278-284.

Nakayama I, Shibazaki M, Yashima-Abo A, Miura G, Sugiyama T, Masuda T and Maesawa C. (2013). Loss of HOXD10 expression induced by upregulation of miR-10b accelerates the migration and invasion activities of ovarian cancer cells. Int J Oncol. 43:63-71.

Napoli C, Grimaldi V, De Pascale MR, Sommese L, Infante T and Soricelli A. (2016). Novel epigenetic-based therapies useful in cardiovascular medicine. World J Cardiol. 8:211-219.

Nayak A, Viale-Bouroncle S, Morsczeck C and Muller S. (2014). The SUMO-specific isopeptidase SENP3 regulates MLL1/MLL2 methyltransferase complexes and controls osteogenic differentiation. Mol Cell. 55:47-58.

Nazarali A, Kim Y and Nirenberg M. (1992). Hox-1.11 and Hox-4.9 homeobox genes. Proc Natl Acad Sci USA. 89:2883-2887.

Nazarali A, Puthucode R, Leung V, Wolf L, Hoa Z and Yeung J. (2000). Temporal and spatial expression of Hoxa-2 during murine palatogenesis. Cell Mol Neurobiol. 20:269-290.

Neer EJ, Schmidt CJ, Nambudripad R and Smith TF. (1994). The ancient regulatory-protein family of WD-repeat proteins. Nature. 371:297-300.

Neer EJ and Smith TF. (2000). A groovy new structure. Proc Natl Acad Sci USA. 97:960-962.

Nicolay DJ, Doucette JR and Nazarali AJ. (2004). Early stages of oligodendrocyte development in the embryonic murine spinal cord proceed normally in the absence of *Hoxa2*. Glia. 48:14-26.

Nielsen BS. (2012). MicroRNA in situ hybridization. Methods Mol Biol. 822:67-84.

Ng SS, Yue WW, Oppermann U and Klose RJ. (2009). Dynamic protein methylation in chromatin biology. Cell Mol Life Sci. 66:407–422.

Nonchev S, Vesque C, Maconochie M, Seitanidou T, Ariza-McNaughton L, Frain M, Marshall H, Sham MH, Krumlauf R and Charnay P. (1996). Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20. Development. 122:543-554.

Noonan EJ, Place RF, Pookot D, Basak S, Whitson JM, Hirata H, Giardina C and Dahiya R. (2009). miR-449a targets HDAC-1 and induces growth arrest in prostate cancer. Oncogene. 28:1714-1724.

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Blüthgen N, Dekker J and Heard E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 485:381-

385.

Nottrott S, Simard MJ and Richter JD. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. Nat Struct Mol Biol. 13:1108–1114.

O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA and Jenuwein T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol. 21:4330–4336.

Odho Z, Southall SM and Wilson JR. (2010). Characterization of a novel WDR5-binding site that recruits RbBP5 through a conserved motif to enhance methylation of histone H3 lysine 4 by mixed lineage leukemia protein-1. J Biol Chem. 285:32967-32976.

Ohhata T, Hoki Y, Sasaki H and Sado T. (2008). Crucial role of antisense transcription across the Xist promoter in Tsix-mediated Xist chromatin modification. Development. 135:227–235.

Okano M, Xie S and Li E. (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nature Genetics. 19:219-220.

Okano M, Bell DW, Haber DA and Li E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 99: 247-257.

Olena AF and Patton JG. (2010). Genomic Organization of microRNAs. J. Cell. Physiol. 222: 540–545.

Özeş AR, Miller DF, Özeş ON, Fang F, Liu Y, Matei D, Huang T and Nephew KP. (2016). NF-κB-HOTAIR axis links DNA damage response, chemoresistance and cellular senescence in ovarian cancer.

Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi, E and Helin K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. 23:4061–4071.

Patel A, Vought VE, Dharmarajan V and Cosgrove MS. (2008). A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. J Biol Chem. 283:32162-32175.

Patel A, Dharmarajan V, Vought VE and Cosgrove MS. (2009). On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. J Biol Chem. 284:24242–24256.

Pearson JC, Lemons D and McGinnis W. (2005). Modulating Hox gene functions during animal body patterning. Nat Rev Genet. 6:893–904.

Penny G, Kay G, Sheardown S, Rastan S and Brockdorff N. (1996). Requirement for Xist in X chromosome inactivation. Nature. 379:131–138.

Pichler G, Wolf P, Schmidt CS, Meilinger D, Schneider K, Frauer C, Fellinger K, Rottach A and

Leonhardt H. (2011). Cooperative DNA and histone binding by Uhrf2 links the two major repressive epigenetic pathways. J Cell Biochem. 112:2585-2593

Philip P, Boija A, Vaid R, Churcher AM, Meyers DJ, Cole PA, Mannervik M and Stenberg P. (2015). CBP binding outside of promoters and enhancers in Drosophila melanogaster. Epigenetics Chromatin. 8:48.

Pillai RS, Bhattacharyya SN, Artus CG, Zoller T, Cougot N, Basyuk E, Bertrand E and Filipowicz W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science. 309:1573–1576.

Pineault KM and Wellik DM. (2014). Hox genes and limb musculoskeletal development. Curr Osteoporos Rep. 12:420-427.

Pinter S, Sadreyev R, Yildirim E, Jeon Y, Ohsumi T, Borowsky M and Lee JT. (2012). Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. Genome Res. 22:1864–1876.

Piunti A and Shilatifard A. (2016). Epigenetic balance of gene expression by Polycomb and COMPASS families. Science. 352:aad9780.

Pogribny IP. (2010). Epigenetic events in tumorigenesis: putting the pieces together. Exp Oncol. 3:132–136.

Prokhortchouk A, Hendrich B, Jorgensen H, Ruzov A, Wilm M, Georgiev G, Bird A, and Prokhortchouk E. (2001). The p120 catenin partner Kaiso is a DNA methylationdependent transcriptional repressor. Genes & Dev. 15: 1613–1618.

Qiao R, He Y, Pan B, Xiao S, Zhang X, Li J, Zhang Z, Hong Y, Xing Y and Ren J. (2015). Understanding the molecular mechanisms of human microtia via a pig model of HOXA1 syndrome. Dis Model Mech. 8:611-622.

Qin W, Wolf P, Liu N, Link S, Smets M, La Mastra F, Forné I, Pichler G, Hörl D, Fellinger K, Spada F, Bonapace IM, Imhof A, Harz H and Leonhardt H. (2015). DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. Cell Res. 25:911-929.

Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A and Tyagi S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nature methods. 5: 877-879.

Rasio D, Schichman SA, Negrini M, Canaani E and Croce CM. (1996). Complete exon structure of the ALL1 gene. Cancer Res. 56:1766-1769.

Rehmsmeier M, Steffen P, Hochsmann M and Giegerich R. (2004). Fast and effective prediction of microRNA/target duplexes. RNA. 10:1507–1517.

Ren YK, Xiao Y, Wan XB, Zhao YZ, Li J, Li Y, Han GS, Chen XB, Zou QY, Wang GC, Lu CM, Xu YC and Wang YC. (2015). Association of long non-coding RNA HOTTIP with progression and prognosis in colorectal cancer. Int J Clin Exp Pathol. 8:11458-11463.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E and Chang HY. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 129:1311-1323.

Rinn JL and Chang HY. (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem. 81:145-166.

Rondelet G, Dal Maso T, Willems L and Wouters J. (2016). Structural basis for recognition of histone H3K36me3 nucleosome by human de novo DNA methyltransferases 3A and 3B. J Struct Biol. 194:357-367.

Ronshaugen M, Biemar F, Piel J, Levine M and Lai EC. (2005). The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. Genes Dev. 19:2947–2952.

Rosin JM, Li W, Cox LL, Rolfe SM, Latorre V, Akiyama JA, Visel A, Kuramoto T, Bobola N, Turner EE and Cox TC. (2016). A distal 594bp ECR specifies Hmx1 expression in pinna and lateral facial morphogenesis and is regulated by Hox-Pbx-Meis. Development. Pii: dev. 133736.

Rossetto D, Avvakumov N and Côté J. (2012). Histone phosphorylation: a chromatin modification involved in diverse nuclear events. Epigenetics. 7:1098-1108.

Roux M, Laforest B, Capecchi M, Bertrand N and Zaffran S. (2015). Hoxb1 regulates proliferation and differentiation of second heart field progenitors in pharyngeal mesoderm and genetically interacts with Hoxa1 during cardiac outflow tract development. Dev Biol. 406:247-258.

Ruddle FH, Bartels JL, Bentley KL, Kappen C, Murtha MT and Pendleton JW. (1994). Evolution of Hox genes. Annu Rev Genet. 28:423-442.

Sado T, Hoki Y and Sasaki H. (2005). Tsix silences Xist through modification of chromatin structure. Dev Cell. 9:159–165.

Saito M and Ishikawa F. (2002). The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2. J Biol Chem. 277: 35434-35439.

Saito Y, Liang G, Egger G, Friedman JM, Chuang JC, Coetzee GA and Jones PA. (2006). Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell. 9:435–443.

Santagati F, Minoux M, Ren SY and Rijli FM. (2005). Temporal requirement of Hoxa2 in cranial neural crest skeletal morphogenesis. Development. 132:4927-4936.

Santini S, Boore JL and Meyer A. (2003). Evolutionary conservation of regulatory elements in vertebrate Hox gene clusters. Genome Res. 13:1111-1122.

Santos KF, Mazzola TN and Carvalho HF. (2005). The prima donna of epigenetics: the regulation of gene expression by DNA methylation. Braz J Med Biol Res. 38:1531-1541.

Sasai N and Defossez PA. (2009). Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. Int. J. Dev. Biol. 53: 323-334.

Sasai N, Nakao M and Defossez PA. (2010). Sequence-specific recognition of methylated DNA by human zinc-finger proteins. Nucleic Acids Res. 38:5015-5022.

Sathi GA, Tsujigiwa H, Ito S, Siar CH, Katase N, Tamamura R, Harada H and Nagatsuka H. (2012). Osteogenic genes related to the canonic WNT pathway are down-regulated in ameloblastoma. Oral Surg Oral Med Oral Pathol Oral Radiol. 114:771-777.

Sato F, Tsuchiya S, Meltzer SJ and Shimizu K. (2011). MicroRNAs and epigenetics. FEBS J. 278:1598-1609.

Sawicka A and Seiser C. (2014). Sensing core histone phosphorylation — A matter of perfect timing. Biochim Biophys Acta. 1839:711-718.

Schmitt AM and Chang HY. (2016). Long Noncoding RNAs in Cancer Pathways. Cancer Cell. 29(4):452-463.

Schmitz KM, Mayer C, Postepska A and Grummt I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNAgenes. Genes Dev. 24:2264-2269.

Schmitz SU, Grote P and Herrmann BG. (2016). Mechanisms of Long Noncoding RNA Function in Development and Disease. Cell Mol Life Sci. 73:2491-2509.

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B and Cavalli G. (2007). Genome Regulation by Polycomb and Trithorax Proteins. Cell. 128:735–745.

Schuettengruber B, Martinez AM, Iovino N, Cavalli G. (2011). Trithorax group proteins: switching genes on and keeping them active. Nat Rev Mol Cell Biol. 12:799-814.

Schuetz A, Allali-Hassani A, Martin F, Loppnau P, Vedadi M, Bochkarev A, Plotnikov AN, Arrowsmith CH and Min J. (2006). Structural basis for molecular recognition and presentation of histone H3 by WDR5. EMBO J. 25:4245-4252.

Scott GK, Mattie MD, Berger CE, Benz SC and Benz CC. (2006). Rapid alteration of microRNA levels by histone deacetylase inhibition. Cancer Res. 66:1277–1281.

Scott MP, Tamkun JW and Hartzell GW. (1989). The structure and function of the homeodomain. Biochim. Biophys. Acta. 989:25-46.

Sehat B, Tofigh A, Lin Y, Trocmé E, Liljedahl U, Lagergren J and Larsson O. (2010). SUMOylation mediates the nuclear translocation and signaling of the IGF-1 receptor. Sci Signal. 3:ra10.

Seifert A, Werheid DF, Knapp SM and Tobiasch E. (2015). Role of Hox genes in stem cell differentiation. World J Stem Cells. 7:583-595.

Shah N and Sukumar S. (2010). The Hox genes and their roles in oncogenesis. Nat Rev Cancer. 10:361-371.

Shah S, Lubeck E, Schwarzkopf M, He TF, Greenbaum A, Sohn CH, Lignell A, Choi HM, Gradinaru V, Pierce NA and Cai L. (2016). Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. Development. 143:2862-2867.

Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-KosekiI Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M and Koseki H. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature. 450:908-912.

Sheikh BN. (2014). Crafting the brain - role of histone acetyltransferases in neural development and disease. Cell Tissue Res. 356:553-573.

Shen WF, Hu YL, Uttarwar L, Passegue E and Largman C. (2008). MicroRNA-126 Regulates HOXA9 by Binding to the Homeobox. Mol. Cell. Biol. 28:4609–4619.

Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T and Sasaki H. (2013). Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet. 9:e1003439.

Shinsky SA, Hu M, Vought VE, Ng SB, Bamshad MJ, Shendure J and Cosgrove MS. (2014). A non-active-site SET domain surface crucial for the interaction of MLL1 and the RbBP5/Ash2L heterodimer within MLL family core complexes. J Mol Biol. 426:2283-2299.

Shinsky SA, Monteith KE, Viggiano S and Cosgrove MS. (2015). Biochemical Reconstitution and Phylogenetic Comparison of Human SET1 Family Core Complexes Involved in Histone Methylation. J Biol Chem. 290:6361-6375.

Shimbo T, Du Y, Grimm SA, Dhasarathy A, Mav D, Shah RR, Shi H and Wade PA. (2013). MBD3 localizes at promoters, gene bodies and enhancers of active genes. PLoS Genet. 9:e1004028.

Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R and Oberdoerffer S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature. 479:74-79.

Simon MD. Wang CI, Kharchenko PV, West JA, Chapman BA, Alekseyenko AA, Borowsky ML, Kuroda MI and Kingston RE. (2011) The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci USA. 108:20497-20502.

Skinner MK. (2011). Role of epigenetics in developmental biology and transgenerational inheritance. Birth Defects Research (Part C) 93:51–55

Smith TF. (2008). Diversity of WD-repeat Proteins. Subcell Biochem. 48:20-30.

Smith TM, Lozanoff S, Iyyanar PP, Nazarali AJ. (2013). Molecular signaling along the anterior-posterior axis of early palate development. Front Physiol. 3:488.

Smith TM, Wang X, Zhang W, Kulyk W and Nazarali AJ.（2009）. Hoxa2 plays a direct role in murine palate development. Developmental dynamics. 238:2364-2373.

Sobko A, Ma H and Firtel RA. (2002). Regulated SUMOylation and ubiquitination of DdMEK1 is required for proper chemotaxis. Dev Cell. 2:745-756.

Song JJ and Kingston RE. (2008). WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. J Biol Chem. 283:35258-35264.

Song RR, Zou L, Zhong R, Zheng XW, Zhu BB, Chen W, Liu L and Miao XP. (2011). An integrated meta-analysis of two variants in HOXA1/HOXB1 and their effect on the risk of autism spectrum disorders. PLoS One. 6:e25603.

Spivakov M and Fisher AG. (2007). Epigenetic signatures of stem-cell identity. Nat Rev Genet. 8:263-271.

Stark A, Brennecke J, Bushati N, Russell RB and Cohen SM. (2005). Animal MicroRNAs Confer Robustness to Gene Expression and Have a Significant Impact on 3'UTR Evolution. Cell. 123:1133-1146.

Steward MM, Lee JS, O'Donovan A, Wyatt M, Bernstein BE and Shilatifard A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nat Struct Mol Biol. 13: 852-854.

Stoicea N, Du A, Lakis DC, Tipton C, Arias-Morales CE, Bergese SD. (2016). The MiRNA Journey from Theory to Practice as a CNS Biomarker. Front Genet. 7: 11.

Studer M, Gavalas A, Marshall H, Ariza-McNaughton L, Rijili FM, Chambon P and Krumlauf R. (1998). Genetic interactions between *Hoxa1* and *Hoxb1* reveal new roles in regulation of early hindbrain patterning. Development. 125:1025-1036.

Suetake I, Shinozaki F, Miyagawa J, Takeshima H and Tajima S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J. Biol. Chem. 279: 27816-27823.

Sun X, Du P, Yuan W, Du Z, Yu M, Yu X and Hu T. (2015). Long non-coding RNA HOTAIR regulates cyclin J via inhibition of microRNA-205 expression in bladder cancer. Cell Death Dis. 6:e1907.

Sun XX, He X, Yin L, Komada M, Sears RC and Dai MS. (2015). The nucleolar ubiquitin-specific protease USP36 deubiquitinates and stabilizes c-Myc. Proc Natl Sci U S A. 112:3734-3739.

Svingen T and Tonissen KF. (2006). Hox transcription factors and their elusive mammalian gene targets. Heredity (Edinb). 97:88-96.

Tan DP, Ferrante J, Nazarali A, Shao X, Kozak CA, Guo V and Nirenberg M. (1992). Murine Hox-1.11 homeobox gene structure and expression. Proc Natl Acad U S A. 89:6280-6284.

Tang G. (2005). siRNA and miRNA: an insight into RISCs. TRENDS in Biochemical Sciences 30:106-114.

Tang L, Shen H, Li X, Li Z, Liu Z, Xu J, Ma S, Zhao X, Bai X, Li M, Wang Q and Ji J. (2016). MiR-125a-5p decreases after long non-coding RNA HOTAIR knockdown to promote cancer cell apoptosis by releasing caspase 2. Cell Death Dis. 7:e2137.

Tang T, Eldabaje R and Yang L. (2016). Current Status of Biological Therapies for the Treatment of Metastatic Melanoma. Anticancer Res. 36:3229-3241.

Tang W, Jiang Y, Mu X, Xu L, Cheng W and Wang X. (2014). MiR-135a functions as a tumor suppressor in epithelial ovarian cancer and regulates HOXA10 expression. Cell Signal. 26:1420-1426.

Tavella S and Bobola N. (2009). Expressing Hoxa2 across the entire endochondral skeleton alters the shape of the skeletal template in a spatially restricted fashion. Differentiation 79:194-202.

Taverna SD, Li H, Ruthenburg AJ, Allis CD and Patel DJ (2007). How chromatin- binding modules interpret histone modifications: lessons from professional pocket pickers. Nat StructMol Biol. 14:1025-1040.

Thompson BA, Tremblay V, Lin G and Bochar DA. (2008). CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. Mol Cell Biol. 28:3894-3904.

Thomsen S, Azzam G, Kaschula R, Williams LS and Alonso CR. (2010). Developmental RNA processing of 3'UTRs in Hox mRNAs as a context-dependent mechanism modulating visibility to microRNAs. Development. 137:2951-2960.

Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, Diaz MO, Scacheri PC and Harte PJ. (2009). CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development. 136:3131-3141.

Tie F, Banerjee R, Saiakhova AR, Howard B, Monteith KE, Scacheri PC, Cosgrove MS and Harte PJ. (2014). Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. Development. 141:1129-1139.

Tijchon E, van Ingen Schenau D, van Opzeeland F, Tirone F, Hoogerbrugge PM, Van Leeuwen FN and Scheijen B. (2015). Targeted Deletion of Btg1 and Btg2 Results in Homeotic Transformation of the Axial Skeleton. PLoS One. 10:e0131481.

Tischfield MA, Bosley TM, Salih MA, Alorainy IA, Sener EC, Nester MJ, Oystreck DT, Chan WM, Andrews C, Erickson RP and Engle EC. (2005). Homozygous HOXA1 mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. Nat. Genet. 37:1035–

1037.

Tremblay V, Zhang P, Chaturvedi CP, Thornton J, Brunzelle JS, Skiniotis G, Shilatifard A, Brand M, Couture JF. (2014). Molecular Basis for DPY-30 Association to COMPASS-like and NURF Complexes. Structure. 22:1821-1830.

Truong K, Lee TD, Li B and Chen Y. (2012). Sumoylation of SAE2 C terminus regulates SAE nuclear localization. J Biol Chem. 287:42611-42619.

Tuddenham L, Wheeler G, Ntounia-Fousara S, Waters J, Hajihosseini MK, Clark I and Dalmay T. (2006). The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. FEBS Lett. 580:4214–4217.

Tyler DM, Okamura K, Chung WJ, Hagen JW, Berezikov E, Hannon GJ and Lai EC. (2008). Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. Genes Dev. 22:26-36.

Unoki M, Nishidate T and Nakamura Y. (2004). ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene. 23: 7601-7610.

Usmani A, Shoro AA, Shirazi B and Memon Z. (2016). Investigative and extrapolative role of microRNAs' genetic expression in breast carcinoma. Pak J Med Sci. 32:766-772.

Uysal F, Akkoyunlu G and Ozturk S. (2015). Dynamic expression of DNA methyltransferases (DNMTs) in oocytes and early embryos. Biochimie. 116:103-113.

van der Wijst MG, Venkiteswaran M, Chen H, Xu GL, Plösch T and Rots MG. (2015). Local chromatin microenvironment determines DNMT activity: from DNA methyltransferase to DNA demethylase or DNA dehydroxymethylase. Epigenetics. 10:671-676.

Valencia K and Wutz A. (2015). Recent insights into the regulation of X-chromosome inactivation. Adv. Genomics Genet. 5:227– 238.

Valeyev NV, Downing AK, Sondek J and Deane C. (2008). Electrostatic and functional analysis of the seven-bladed WD beta-propellers. Evol Bioinform Online. 4:203-216.

Valinezhad Orang A, Safaralizadeh R and Kazemzadeh-Bavili M. (2014). Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation. Int J Genomics. 2014:970607.

Valor LM, Viosca J, Lopez-Atalaya JP and Barco A. (2013). Lysine acetyltransferases CBP and p300 as therapeutic targets in cognitive and neurodegenerative disorders. Curr Pharm Des. 19:5051-5064.

Voncken JW, Roelen BA, Roefs M, de Vries S, Verhoeven E, Marino S, Deschamps J, and van Lohuizen M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc Natl Acad Sci U S A. 100: 2468–2473.

Wallace RG, Twomey LC, Custaud MA, Moyna N, Cummins PM, Mangone M, Murphy RP. (2016). Potential Diagnostic and Prognostic Biomarkers of Epigenetic Drift within the Cardiovascular Compartment. Biomed Res Int. 2016:2465763.

Wall MA, Coleman DE, Lee E, Iniguez-Lluhi JA, PosnerBA, Gilman AG and Sprang SR. (1995). The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. Cell. 83:1047-1058.

Walsh CP and Xu GL. (2006). Cytosine methylation and DNA repair. Curr Top Microbiol Immunol. 301: 283-315.

Wang C, Li G, Wu Y, Xi J and Kang J. (2016). LincRNA1230 inhibits the differentiation of mouse ES cells towards neural progenitors. Sci China Life Sci. 59:443-454.

Wang F, Marshall CB and Ikura M. (2013). Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. Cell Mol Life Sci. 70:3989-4008.

Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature. 472:120-124.

Wang M, Doucette JR and Nazarali AJ. (2011). Conditional Tet-Regulated Over-Expression of Hoxa2 in CG4 Cells Increases Their Proliferation and Delays Their Differentiation into Oligodendrocyte-like Cells Expressing Myelin Basic Protein. Cell Mol Neurobiol. 31:875–886.

Wang X. (2013). Mechanism of Valproic Acid Induced Dysmorphogenesis Via Oxidative Stress and Epigenetic Regulation at the Hoxa2 Gene Promoter. PhD Thesis, University of Saskatchewan.

Wang YY, Liu LJ, Zhong B, Liu TT, Li Y, Yang Y, Ran Y, Li S, Tien P and Shu HB. (2010). WDR5 is essential for assembly of the VISA-associated signaling complex and virus-triggered IRF3 and NF-κB activation. Proc Natl Acad Sci U S A. 107:815–820.

Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M and Schübeler D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet. 39:457–466.

Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, Waldvogel B, Vannier C, Darling D, zur Hausen A, Brunton VG, Morton J, Sansom O, Schüler J, Stemmler MP, Herzberger C, Hopt U, Keck T, Brabletz S and Brabletz T. (2009). The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. Nat Cell Biol. 11:1487–1495.

Woltering JM and Durston AJ. (2008). MiR-10 represses HoxB1a and HoxB3a in zebrafish. PLoS ONE. 3:e1396.

Wong SF, Agarwal V, Mansfield JH, Denans N, Schwartz MG, Prosser HM, Pourquié O, Bartel DP, Tabin CJ and McGlinn E. (2015). Independent regulation of vertebral number and vertebral identity by microRNA-196 paralogs. Proc Natl Acad Sci USA. 112:E4884-4893.

Wood KH and Zhou Z. (2016). Emerging Molecular and Biological Functions of MBD2, a Reader of DNA Methylation. Front Genet. 7:93.

Wood LD, Irvin BJ, Nucifora G, Luce KS and Hiebert SW. (2003). Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor. Proc Natl Acad Sci U S A. 100:3257-3262.

Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH and Allis CD. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell. 121:859-872.

Xiao F, Bai Y, Chen Z, Li Y, Luo L, Huang J, Yang J, Liao H and Guo L. (2014). Downregulation of HOXA1 gene affects small cell lung cancer cell survival and chemoresistance under the regulation of miR-100. Eur J Cancer. 50:1541-1554.

Xie M, Sun M, Zhu YN, Xia R, Liu YW, Ding J, Ma HW, He XZ, Zhang ZH, Liu ZJ, Liu XH and De W. (2015). Long noncoding RNA HOXA-AS2 promotes gastric cancer proliferation by epigenetically silencing P21/PLK3/DDIT3 expression. Oncotarget. 6:33587-33601.

Xu L, Chen YC, Nakajima S, Chong J, Wang L, Lan L, Zhang C and Wang D. (2014). A chemical probe targets DNA 5-formylcytosine sites and inhibits TDG excision, polymerases bypass, and gene expression. Chem Sci. 5:567-574.

Yan S and Jiao K. (2016). Functions of miRNAs during Mammalian Heart Development. Int J Mol Sci. 17:pii:E789.

Yang C, Wang Y, Zhang F, Sun G, Li C, Jing S, Liu Q and Cheng Y. (2013). Inhibiting UHRF1 expression enhances radiosensitivity in human esophageal squamous cell carcinoma. Mol Biol Rep. 40:5225-5235.

Yang X, Han H, De Carvalho DD, Lay FD, Jones PA and Liang G. (2014). Gene Body Methylation Can Alter Gene Expression and Is a Therapeutic Target in Cancer. Cancer Cell. 26:577-590.

Yang Y, Kucukkal TG, Li J, Alexov E and Cao W. (2016). Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. ACS Chem Biol. 11:2706-2715.

Yang YW, Flynn RA, Chen Y, Qu K, Wan B, Wang KC, Lei M and Chang HY. (2014). Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. Elife. 3:e02046.

Yao B, Christian KM, He C, Jin P, Ming GL and Song H. (2016). Epigenetic mechanisms in neurogenesis. Nat Rev Neurosci. 17:537-549.

Yekta S, Shih IH and Bartel DP. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science. 304:594–596.

Yekta S, Tabin CJ and Bartel DP. (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. Nature Review. 9:789-796.

Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS, Weng Z, Rando OJ and Fazzio TG. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell 147:1498–1510.

Yu H, Lindsay J, Feng ZP, Frankenberg S, Hu Y, Carone D, Shaw G, Pask AJ, O'Neill R, Papenfuss AT and Renfree MB. (2012). Evolution of coding and non-coding genes in HOX clusters of a marsupial. BMC Genomics. 13:251.

Yu SL, Lee DC, Sohn HA, Lee SY, Jeon HS, Lee JH, Park CG, Lee HY, Yeom YI, Son JW, Yoon YS and Kang J. (2015). Homeobox A9 directly targeted by miR-196b regulates aggressiveness through nuclear Factor-kappa B activity in non-small cell lung cancer cells. Mol Carcinog. doi: 10.1002/mc.22439.

Yue M, Charles Richard JL and Ogawa Y. (2015). Dynamic interplay and function of multiple noncoding genes governing X chromosome inactivation. Biochim Biophys Acta. 1859:112-120.

Zhang P, Bergamin E and Couture JF. (2013). The many facets of MLL1 regulation. Biopolymers. 99:136-145.

Zhang S, Tang B, Fan C, Shi L, Zhang X, Sun L and Li Z. (2015). Effect of DNMT inhibitor on bovine parthenogenetic embryo development. Biochem Biophys Res Connun. 466:505-511.

Zhang SR, Yang JK, Xie JK and Zhao LC. (2016). Long noncoding RNA HOTTIP contributes to the progression of prostate cancer by regulating HOXA13. Cell Mol Biol (Noisy-le-grand). 62:84-88.

Zhang W. (2003). Temporal and spatial expression of Hoxa2 gene in the developing mouse palate and the effects of valporic acid on Hoxa2 expression during murine palatogenesis. M.Sc. Thesis, University of Saskatchewan.

Zhang X, Lian Z, Padden C, Gerstein MB, Rozowsky J, Snyder M, Gingeras TR, Kapranov P, Weissman SM and Newburger PE. (2009). A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood. 113:2526-2534.

Zhang X, Weissman SM and Newburger PE. (2014). Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB4 human promyelocytic leukemia cells. RNA Biol. 11:777-787.

Nguyen AT and Zhang Y. (2011). The diverse functions of Dot1 and H3K79 methylation. Genes Dev. 25:1345-1358.

Zhao H, Zhang X, Frazão JB, Condino-Neto A and Newburger PE. (2013). HOX antisense

lincRNA HOXA-AS2 is an apoptosis repressor in all trans retinoic acid treated NB4 promyelocytic leukemia cells. J Cell Biochem. 114:2375-2383.

Zhao J, Sun BK, Erwin JA, Song JJ and Lee JT. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 322:750-756.

Zhao X, Su J, Wang F, Liu D, Ding J, Yang Y, Conaway JW, Conaway RC, Cao L, Wu D, Wu M, Cai Y and Jin J. (2013). Crosstalk between NSL histone acetyltransferase and MLL/SET complexes: NSL complex functions in promoting histone H3K4 di-methylation activity by MLL/SET complexes. PLoS Genet. 9:e1003940.

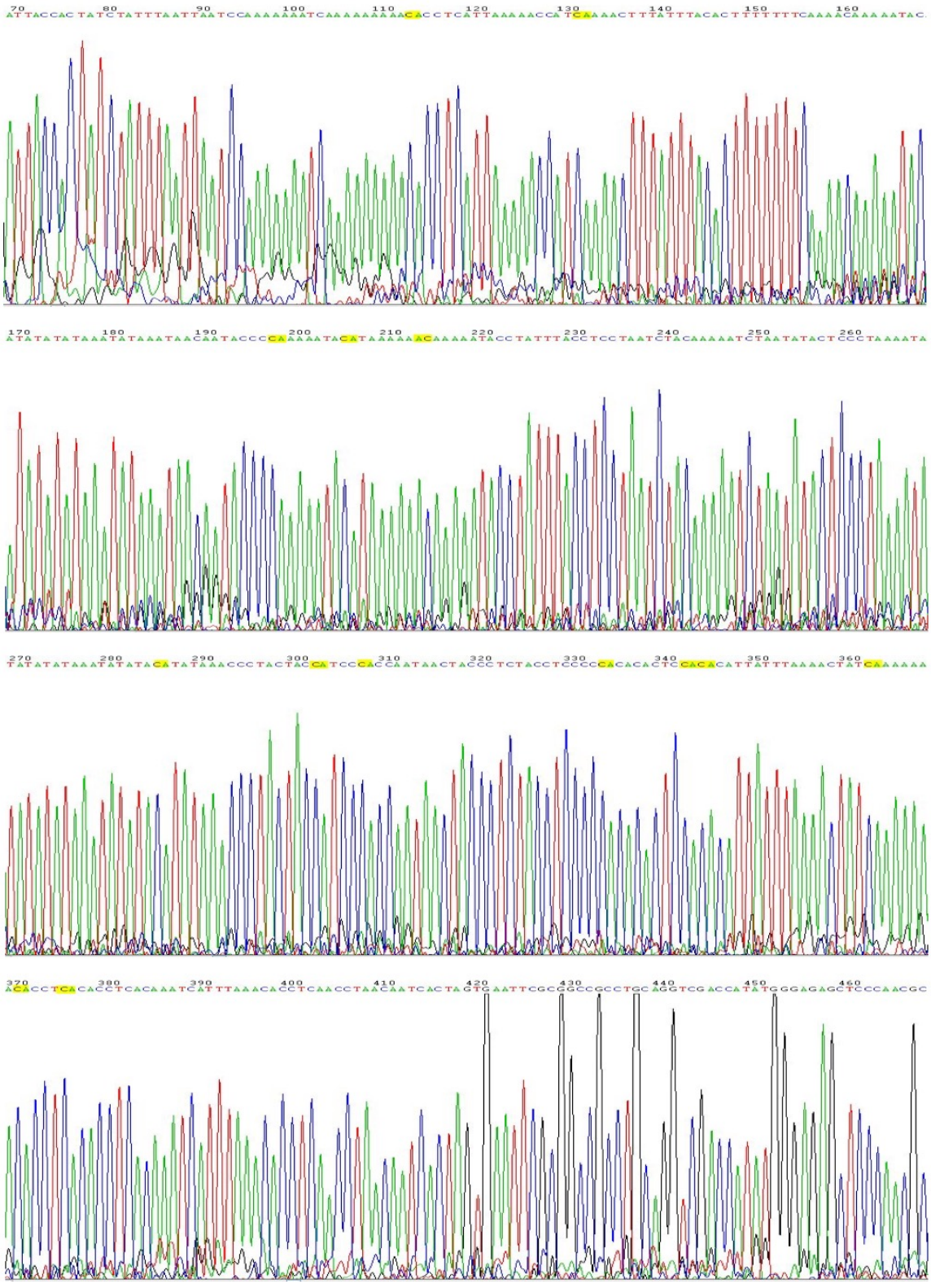
Zhong J, Agha G and Baccarelli AA. (2016). The Role of DNA Methylation in Cardiovascular Risk and Disease: Methodological Aspects, Study Design, and Data Analysis for Epidemiological Studies. Circ Res. 118:119-131.

Zhou X, Ruan J, Wang G and Zhang W. (2007). Characterization and identification of microRNA core promoters in four model species. PLoS Comput Biol. 3:e37.

Zhu ED, Demay MB and Gori F. (2008). Wdr5 is essential for osteoblast differentiation. J Biol Chem. 283:7361-7367.

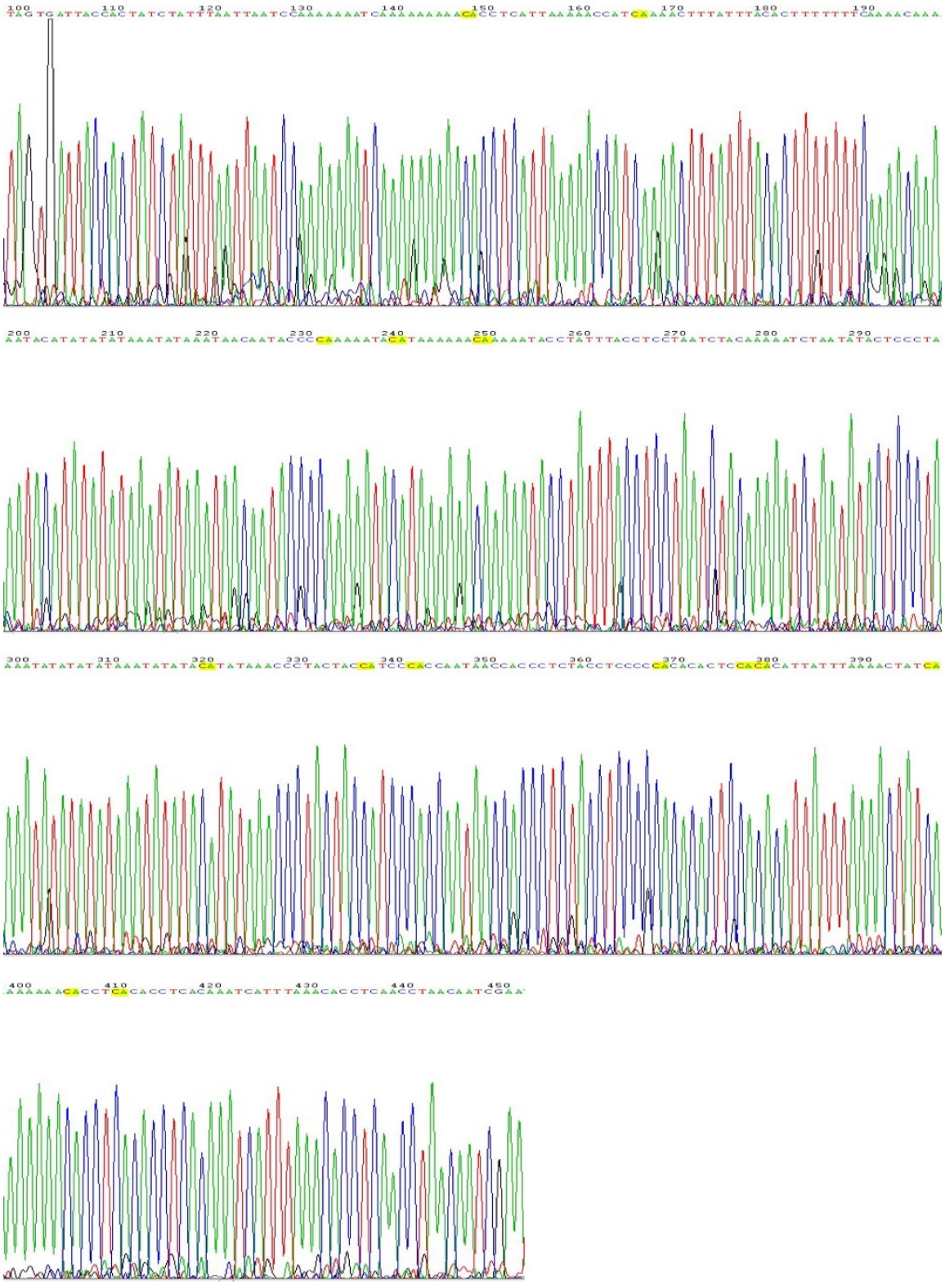
Zhu S, Zhu ED, Provot S and Gori F. (2010). Wdr5 is required for chick skeletal development. J Bone Miner Res. 25:2504-2514.

E12

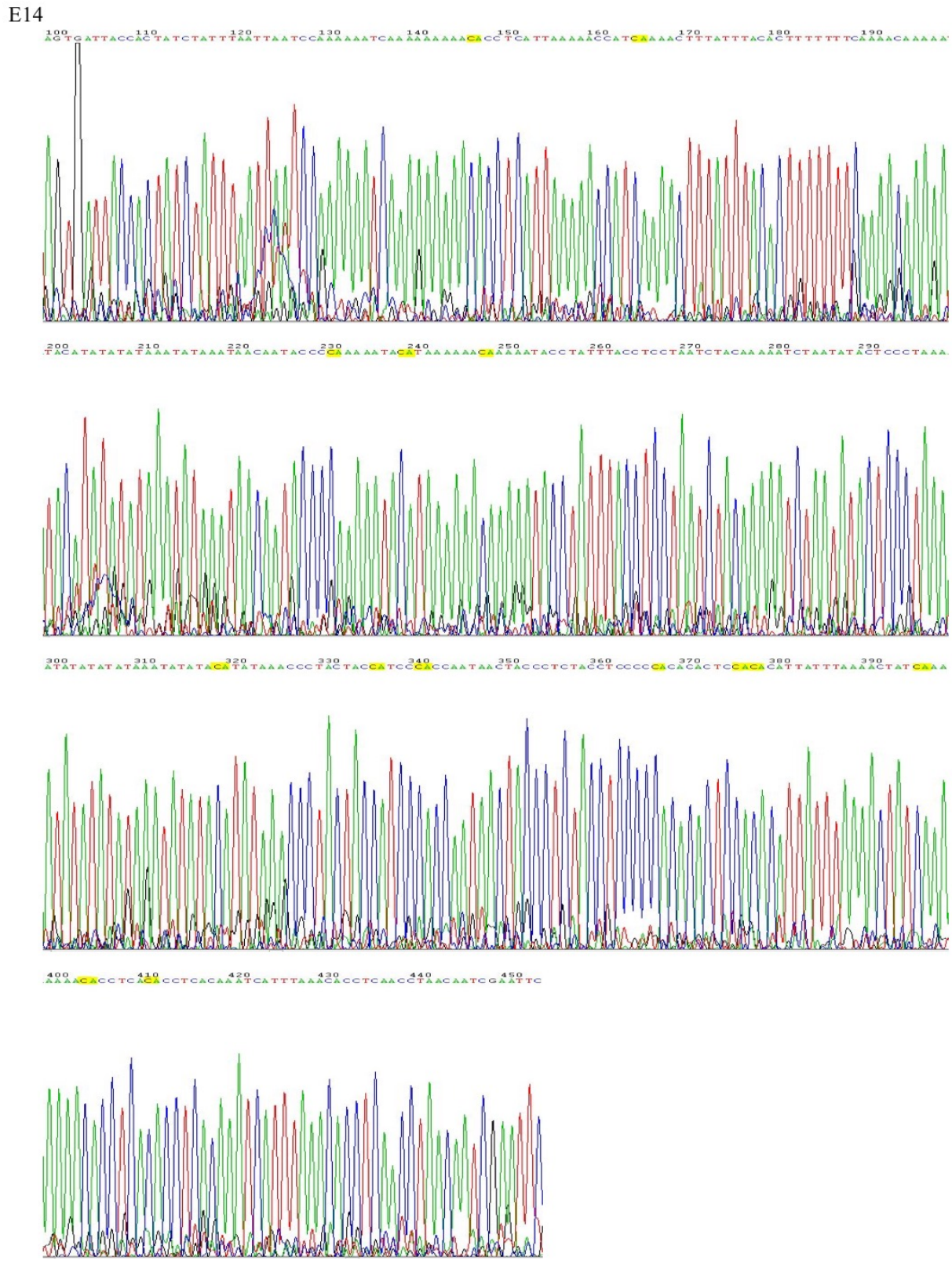


Appendix 1. DNA sequence chromatogram of *Hoxa2* promoter CpG island 1 from E12 mouse palate genomic DNA

E13

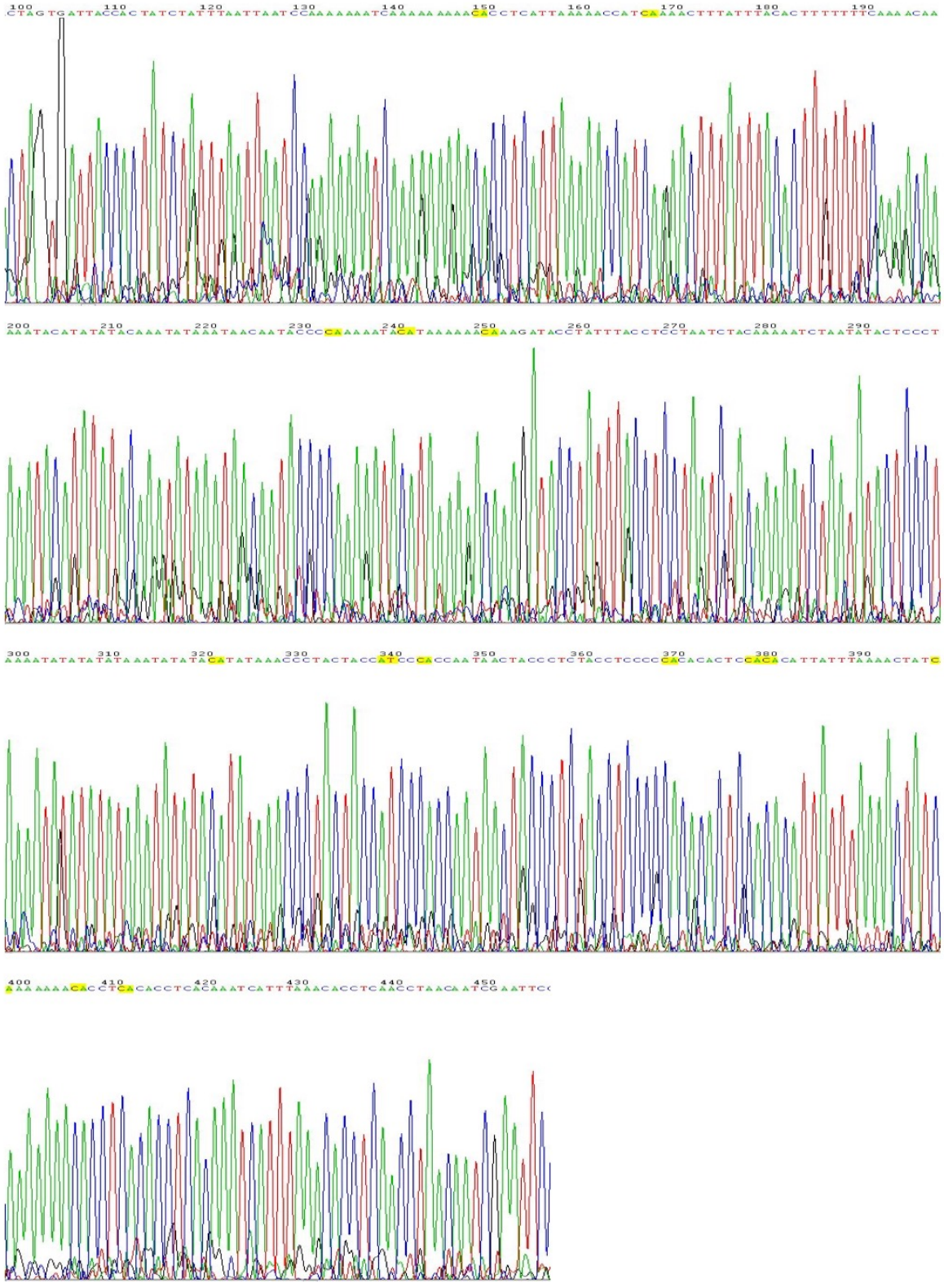


Appendix 2. DNA sequence chromatogram of *Hoxa2* promoter CpG island 1 from E13 mouse palate genomic DNA



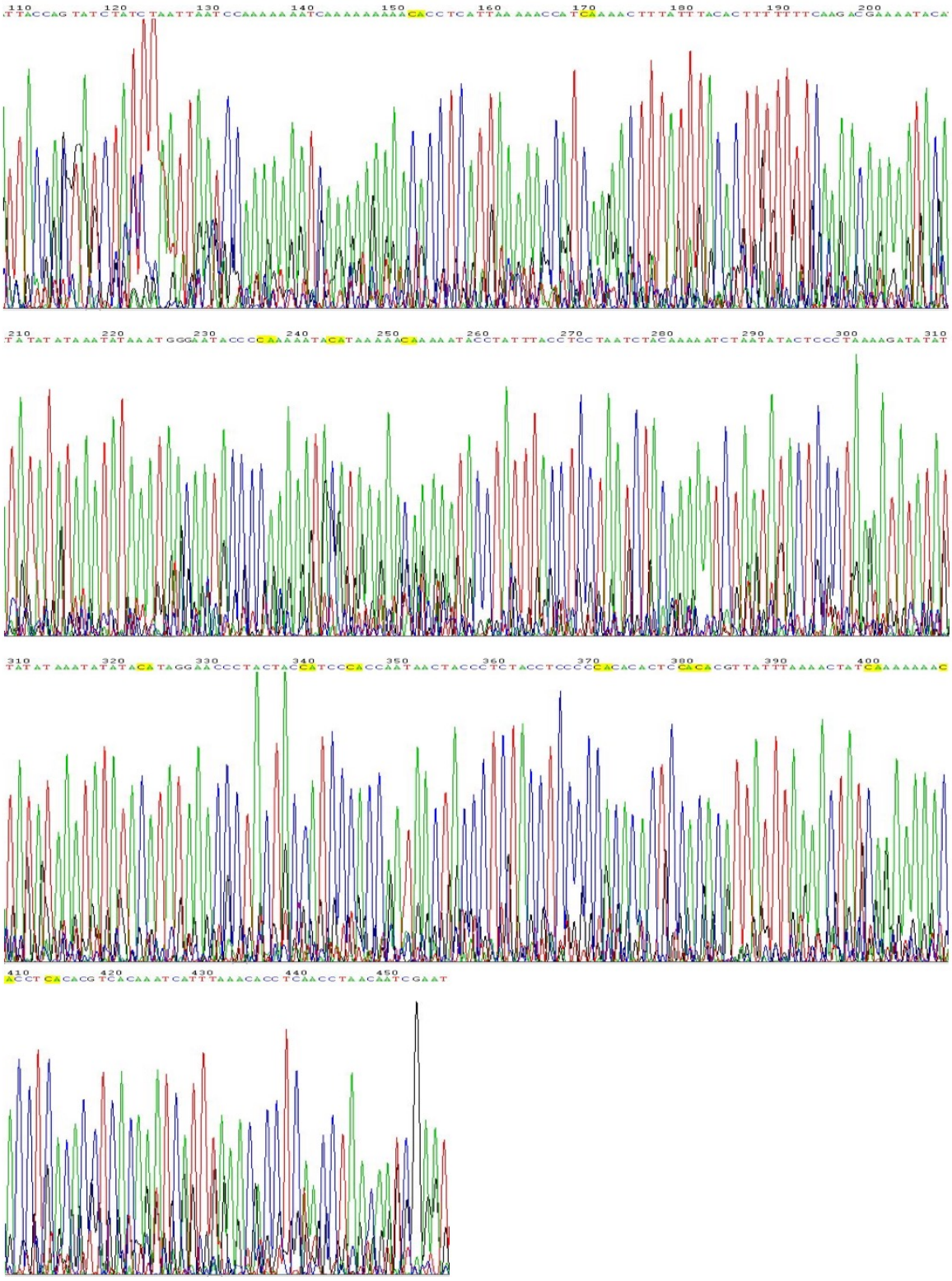
Appendix 3. DNA sequence chromatogram of *Hoxa2* promoter CpG island 1 from E14 mouse palate genomic DNA

E15



Appendix 4. DNA sequence chromatogram of *Hoxa2* promoter CpG island 1 from E15 mouse palate genomic DNA

NIH 3T3

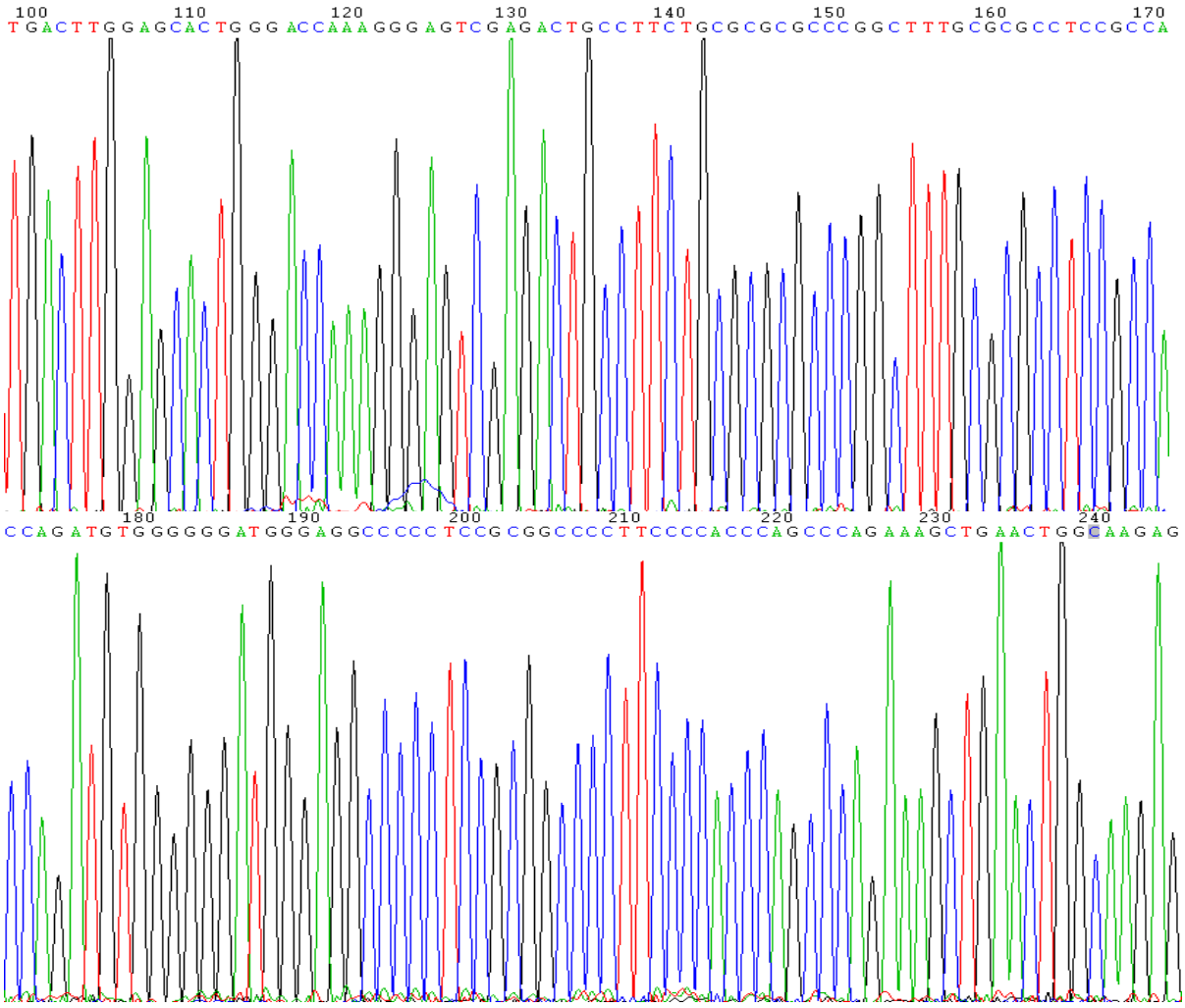


Appendix 5. DNA sequence chromatogram of *Hoxa2* promoter CpG island 1 from NIH 3T3 cell line genomic DNA

EG7

Appendix 6. DNA sequence chromatogram of *Hoxa2* promoter CpG island 1 from EG7 cell line genomic DNA

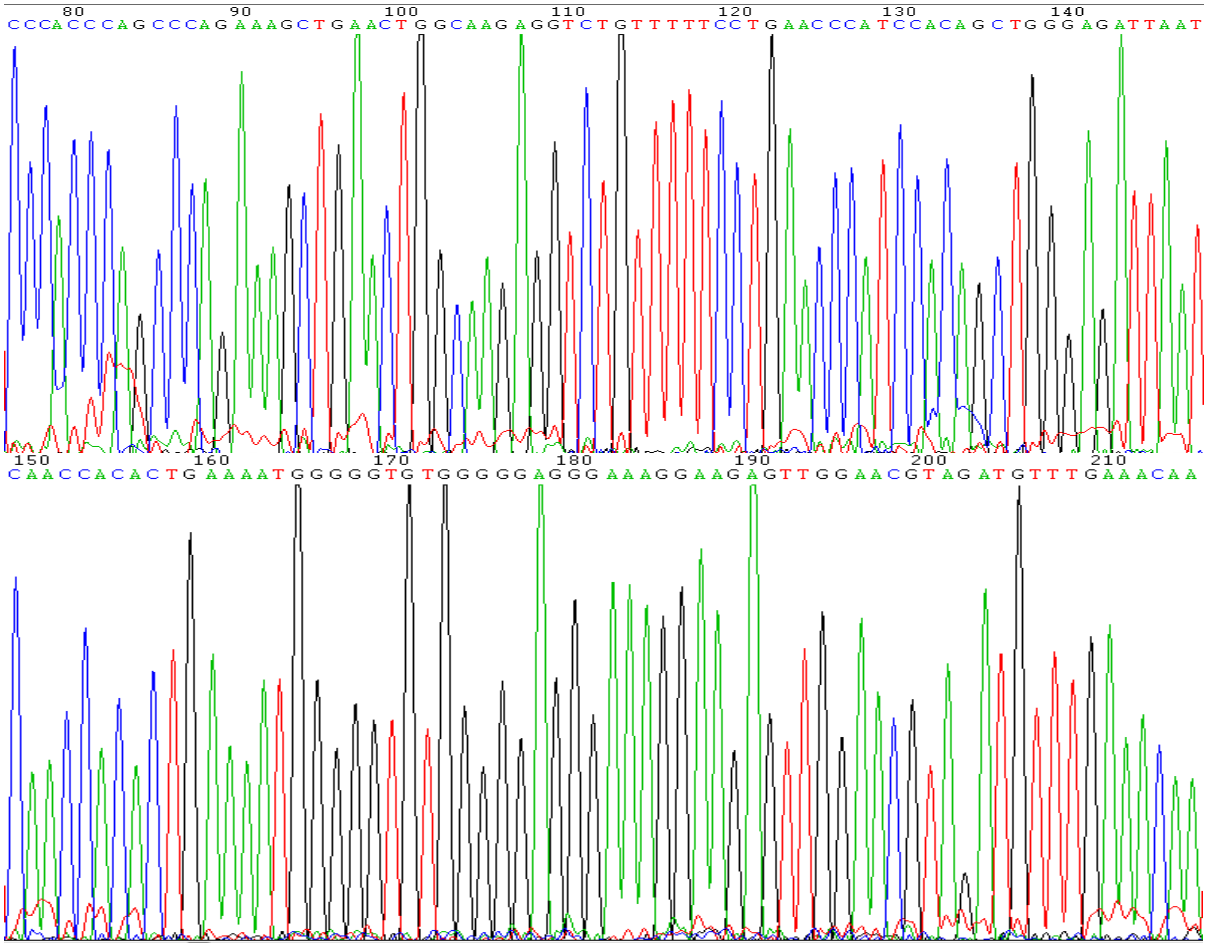
Mush 147-F: TGACTTGGAGCACTGGGA
Mush147-R: CTCTTGCCAGTTCAGCTTTCT



Mush 147 (17-163 bp)
TGACTTGGAGCACTGGGACCAAAGGGAGTCGAGACTGCCTTCTGCGCGCGCCCG
GCTTTGCGCGCCTCCGCCACCAGATGTGGGGGGATGGGAGGCCCCCTCCGCGGCC
CCTTCCCCACCCAGCCCAGAAAGCTGAACTGGCAAGAG

Appendix 7. DNA sequence chromatogram of cDNA fragments amplified with mHotairm1 primer Mush 147 F+R

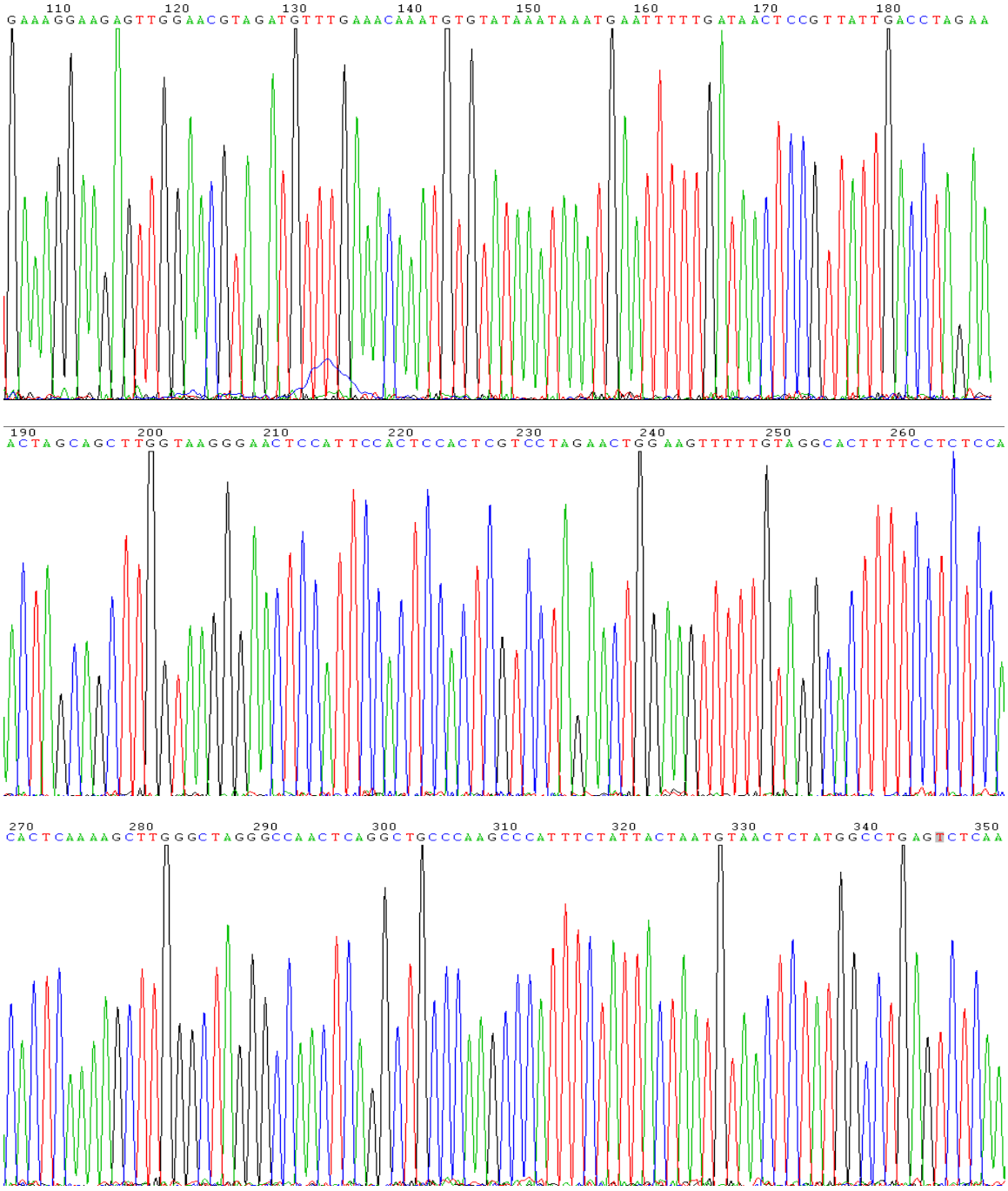
Mush138-F: CCCACCCAGCCCAGAAAG
Mush138-R: GTTTCAAACATCTACGTTCC



Mush 138 (131-268 bp)
CCCCACCCAGCCCAGAAAGCTGAACTGGCAAGAGGTCTGTTTTTCCTGAACCCAT
CCACAGCTGGGAGATTAATCAACCACACTGAAAATGGGGGTGTGGGGGAGGGAA
AGGAAGAGTTGGAACGTAGATGTTTGAAAC

Appendix 8. DNA sequence chromatogram of cDNA fragments amplified with mHotairm1 primer Mush 138 F+R

Mush245-F: GAAAGGAAGAGTTGGAACGTAGA
Mush245-R: TGAGACTCAGGCCATAGAGTTA

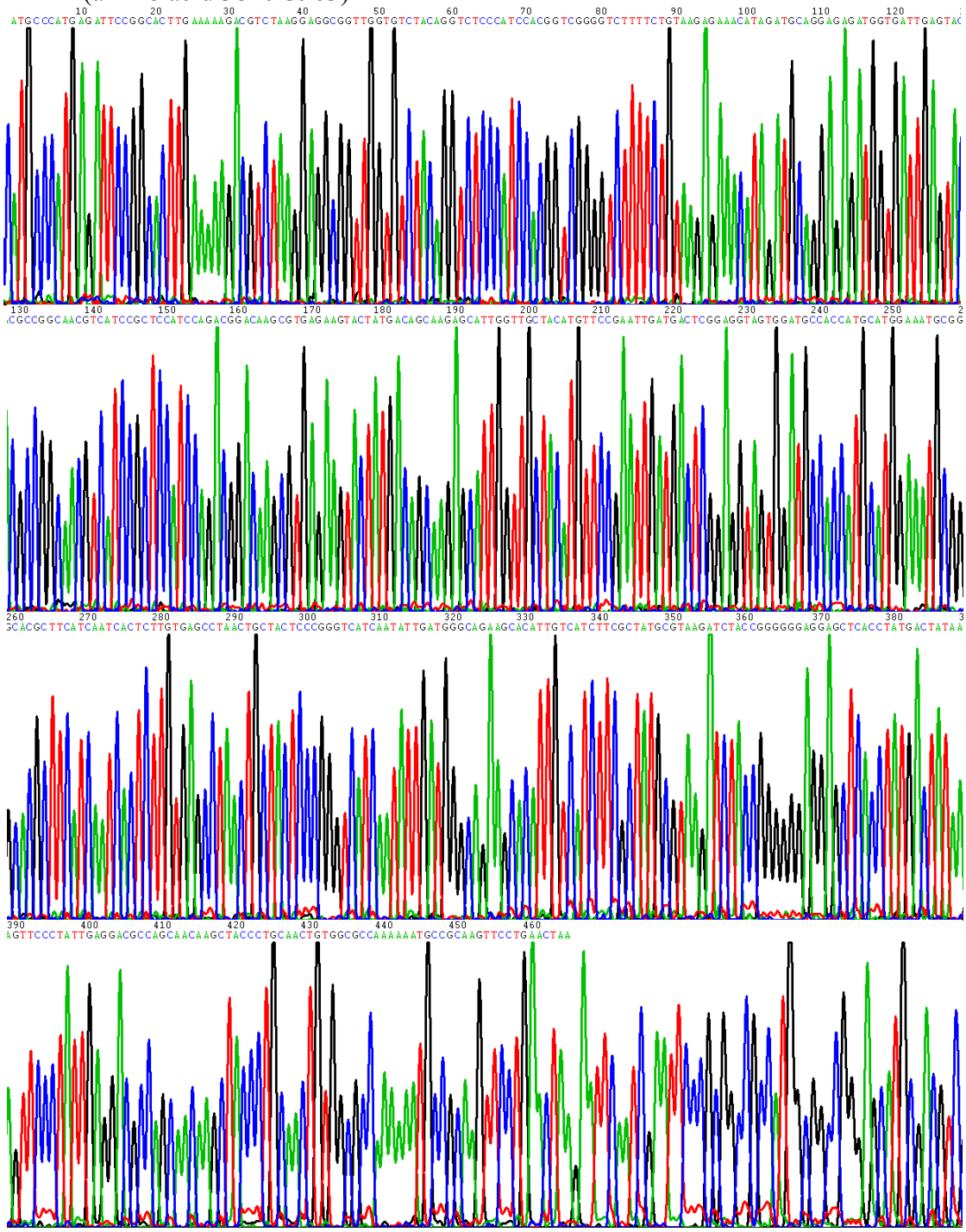


Mush 245 (236-478 bp)

GAAAGGAAGAGTTGGAACGTAGATGTTTGAAACAAATGTGTATAAATAAATGAATT
TTTGATAACTCCGTTATTGACCTAGAAACTAGCAGCTTGGTAAGGGAACTCCATTCC
ACTCCACTCGTCCTAGAACTGGAAGTTTTTGTAGGCACTTTTCCTCTCCACACTCAAA
AGCTTGGGCTAGGGCCAACTCAGGCTGCCCAAGCCCATTTCTATTACTAATGTAACT
CTATGGCCTGAGTCT

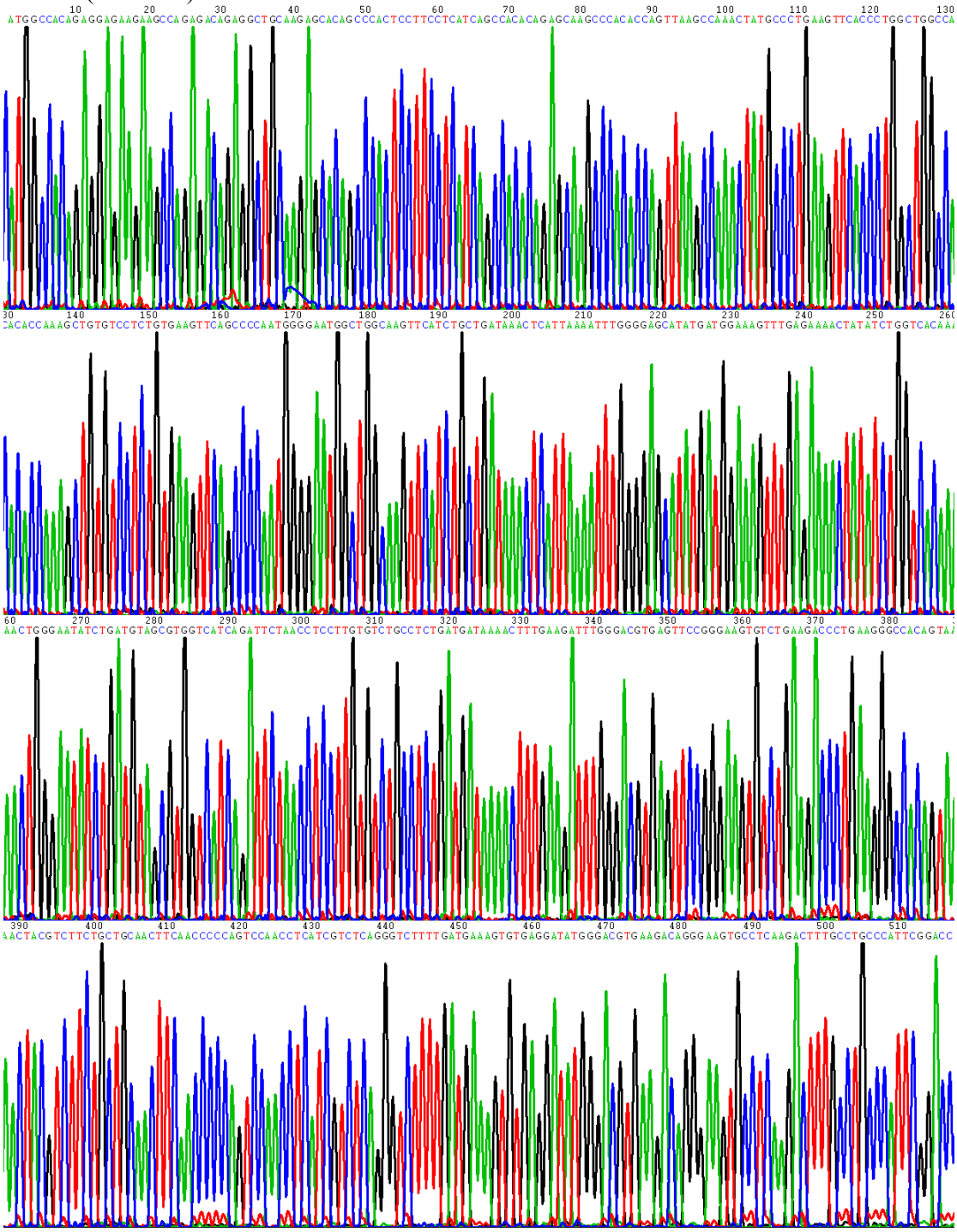
Appendix 9. DNA sequence chromatogram of cDNA fragments amplified with mHotairm1 primer Mush 245 F+R

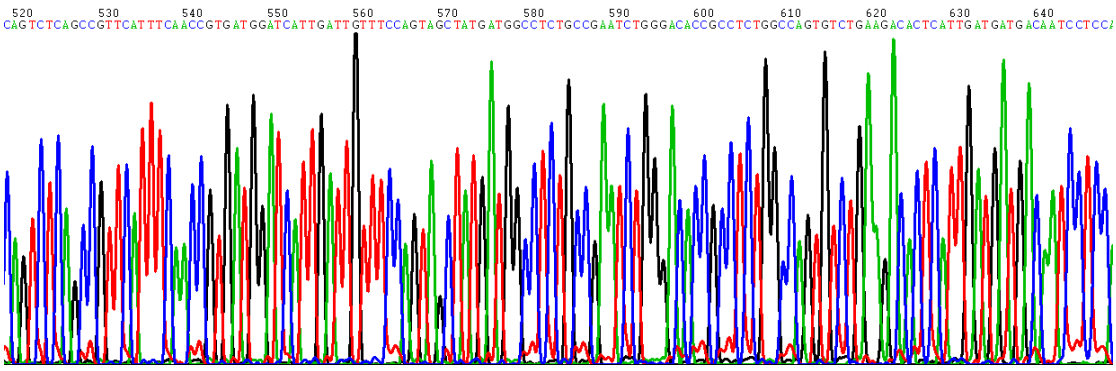
MLL1 (amino acid 3810-3963)



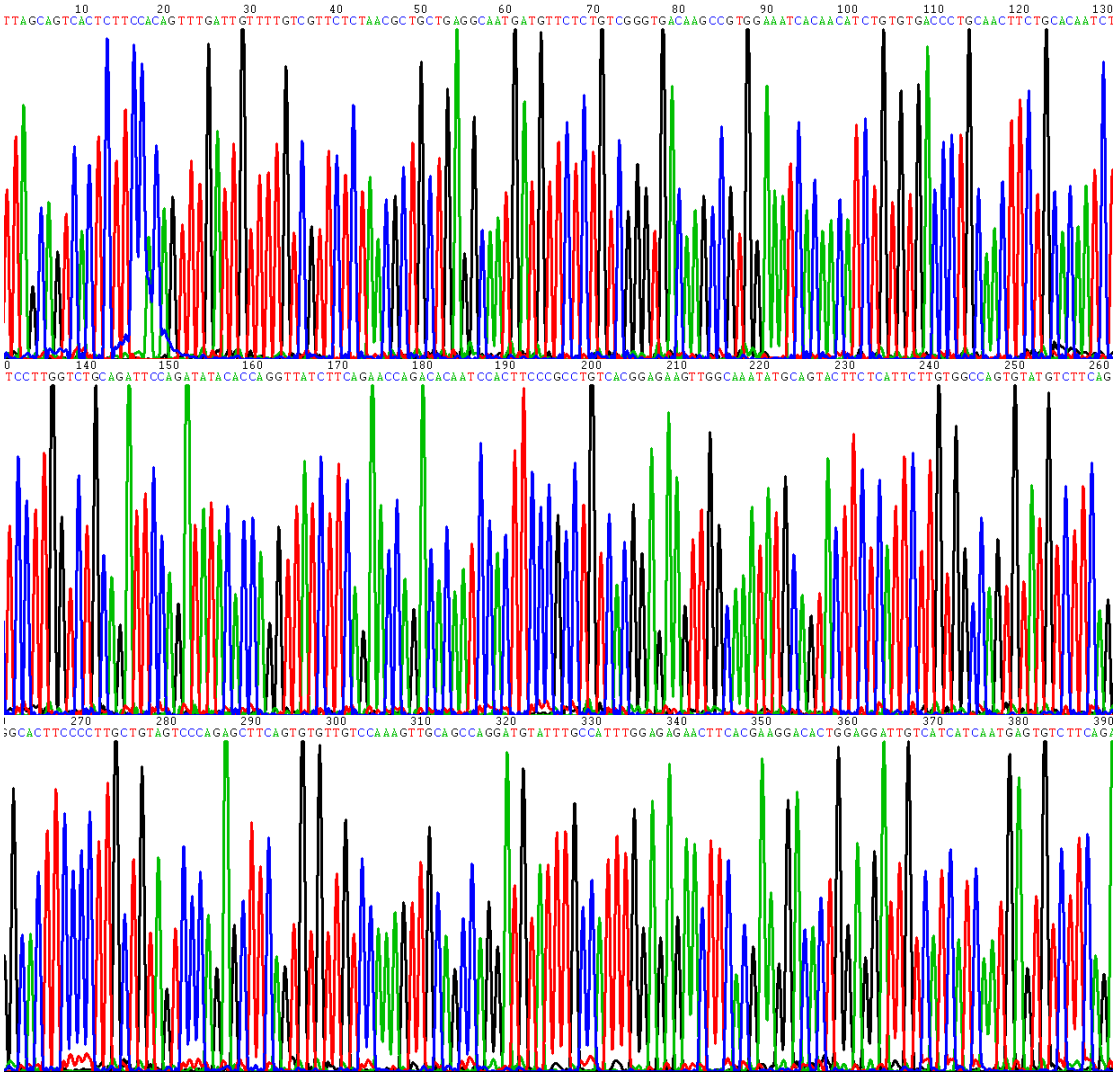
Appendix 10. DNA sequence chromatogram of MLL1 (amino acid 3810-3963) cloning.

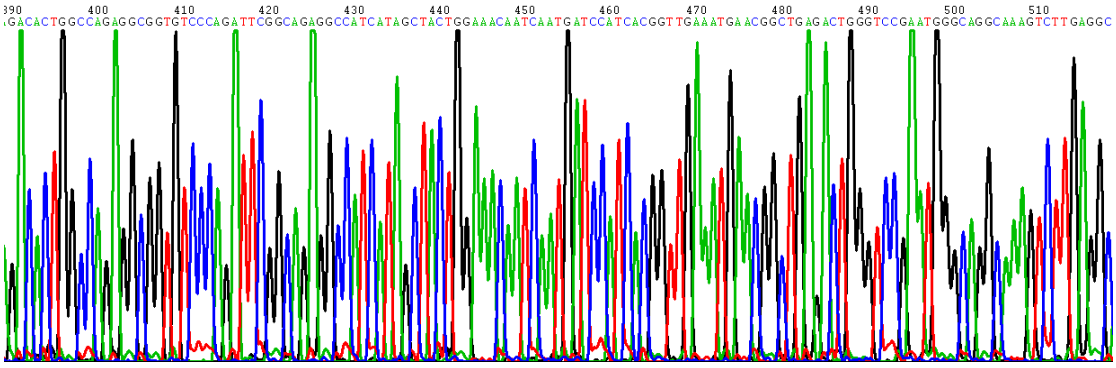
WDR5 (1-640 nt)



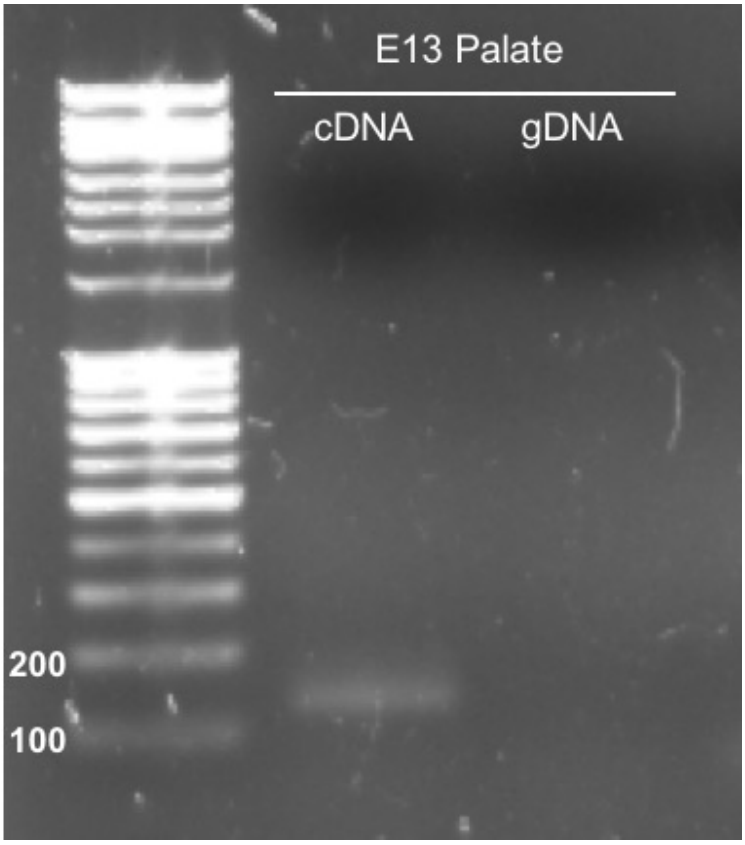


WDR5 (495-1005 nt)





Appendix 11. DNA sequence chromatogram of full length WDR5 cloning.



Appendix 12. *mHotairm1* can only be amplified from cDNA samples using much138 F+R primers. cDNA samples were reverse transcribed from total RNA samples collected from E13 mouse palate. Genomic DNA (gDNA) samples were also collected from E13 mouse palate. mHotairm1 specific primer much138 F+R were used in PCR. A 138bp DNA fragment can only be amplified from cDNA samples but not from genomic DNA.

Copyright Clearance Center

Confirmation Number: 11603725
Order Date: 11/07/2016

Customer Information

Customer: Ran Bi
Account Number: 3001071843
Organization: Ran Bi
Email: rab501@mail.usask.ca
Phone: +1 (306) 966-6357
Payment Method: Credit Card ending in 7710

This is not an invoice

Order Details

Development

Billing Status:
Charged to Credit Card

Order detail ID:	70155824
ISSN:	1477-9129
Publication Type:	e-Journal
Volume:	
Issue:	
Start page:	
Publisher:	COMPANY OF BIOLOGISTS,
Author/Editor:	Company of Biologists

Permission Status: Granted

Permission type: Republish or display content
Type of use: Republish in a thesis/dissertation
Order License Id: 3983640171260

Requestor type	Academic institution
Format	Print
Portion	image/photo
Number of images/photos requested	1
Title or numeric reference of the portion(s)	Figure 3. Hox gene requirement for arch patterning and skeletogenic cranial NCC ground pattern.
Title of the article or chapter the portion is from	Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development
Editor of portion(s)	N/A
Author of portion(s)	Maryline Minoux, Filippo M. Rijli
Volume of serial or monograph	vol. 137 no. 16,
Page range of portion	2611
Publication date of portion	July 27, 2010
Rights for	Main product
Duration of use	Current edition and up to 5 years
Creation of copies for the disabled	no
With minor editing privileges	no

For distribution to	Canada
In the following language(s)	Original language of publication
With incidental promotional use	no
Lifetime unit quantity of new product	Up to 499
Made available in the following markets	education
The requesting person/organization	Ran Bi, University of Saskatchewan
Order reference number	
Author/Editor	Ran Bi
The standard identifier of New Work	N/A
Title of New Work	EPIGENETIC REGULATION OF Hoxa1 AND Hoxa2
Publisher of New Work	University of Saskatchewan
Expected publication date	Jan 2017
Estimated size (pages)	250

Payment Method: CC ending in 7710

Note: This item was charged to your credit card through our **RightsLink service.** More info **$ 3.50**

Total order items: 1 **Order Total: $3.50**

Copyright 2016 Copyright Clearance Center

SPRINGER LICENSE
TERMS AND CONDITIONS

Nov 15, 2016

This Agreement between Ran Bi ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	3966411406766
License date	Oct 12, 2016
Licensed Content Publisher	Springer
Licensed Content Publication	Cellular and Molecular Neurobiology
Licensed Content Title	Hox Genes and Their Candidate Downstream Targets in the Developing Central Nervous System
Licensed Content Author	Z. N. Akin
Licensed Content Date	Jan 1, 2005
Licensed Content Volume Number	25
Licensed Content Issue Number	3
Type of Use	Thesis/Dissertation
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	1
Author of this Springer article	No
Order reference number	
Original figure numbers	figure 1
Title of your thesis / dissertation	EPIGENETIC REGULATION OF Hoxa1 AND Hoxa2
Expected completion date	Jan 2017
Estimated size(pages)	250
Requestor Location	Ran Bi 42 Leddy Cres. Saskatoon, SK S7H 3Y6 Canada Attn: Ran Bi
Billing Type	Invoice
Billing Address	Ran Bi 42 Leddy Cres. Saskatoon, SK S7H 3Y6 Canada

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Nov 15, 2016

This Agreement between Ran Bi ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	3987550993429
License date	Nov 14, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Structural and Molecular Biology
Licensed Content Title	Molecular recognition of histone H3 by the WD40 protein WDR5
Licensed Content Author	Jean-François Couture, Evys Collazo, Raymond C Trievel
Licensed Content Date	Jul 9, 2006
Licensed Content Volume Number	13
Licensed Content Issue Number	8
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1. Crystal structure of WDR5 bound to the N terminus of histone H3
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	EPIGENETIC REGULATION OF Hoxa1 AND Hoxa2
Expected completion date	Jan 2017
Estimated size (number of pages)	250
Requestor Location	Ran Bi College of Pharmacy and Nutrition University of Saskatchewan Saskatoon, SK S7N 5E5 Canada Attn: Ran Bi
Billing Type	Invoice

From: **PNAS Permissions** PNASPermissions@nas.edu
Subject: RE: Figure reprint request
Date: November 14, 2016 at 10:09 AM
To: Bi, Ran rab501@mail.usask.ca

PP

Permission is granted for your use of the figure as described in your message. Please cite the PNAS article in full when re-using the material. Because this material published after 2008, a copyright note is not needed. There is no charge for this material, either. Let us know if you have any questions.

Best regards,
Kay McLaughlin for
Diane Sullenberger
Executive Editor
PNAS

From: Bi, Ran [mailto:rab501@mail.usask.ca]
Sent: Monday, November 14, 2016 7:55 AM
To: PNAS Permissions
Subject: Figure reprint request

Dear Mr./Ms.:

I am a PhD student in University of Saskatchewan and I would like to ask for permission to use a figure in the following publication in my PhD thesis. Please see following for details.
Thanks in advance for your help!

Ran Bi

. Your full name, affiliation, and title: Ran Bi
. Your complete mailing address: College of Pharmacy and Nutrition, University of Saskatchewan, Saskatoon, SK. Canada. S7N 5E5
. phone number: 306-966-6357
. email: rab501@mail.usask.ca
. PNAS volume number, issue number, and issue date: 2011 Dec 20;108(51):20497-20502.
. PNAS article title: The genomic binding sites of a noncoding RNA.
. PNAS authors' names: Simon MD, Wang CI, Kharchenko PV, West JA, Chapman BA, Alekseyenko AA, Borowsky ML, Kuroda MI, Kingston RE.
. Page numbers of items to be reprinted: p20498
. Figure/table number or portion of text to be reprinted: Figure 1.

Requests must also include the following information about the intended use of the material:

. Title of work in which PNAS material will appear: Epigenetic regulation of Hoxa1 and Hoxa2
. Authors/editors of work: Ran Bi
. Publisher of work: University of Saskatchewan
. Retail price of work: N/A
. Number of copies of work to be produced: N/A
. Intended audience: N/A
. Whether work is for nonprofit or commercial use: Nonprofit use. Use in PhD thesis.

11200 Rockville Pike
Suite 302
Rockville, Maryland 20852

August 19, 2011

American Society for Biochemistry and Molecular Biology

To whom it may concern,

It is the policy of the American Society for Biochemistry and Molecular Biology to allow reuse of any material published in its journals (the Journal of Biological Chemistry, Molecular & Cellular Proteomics and the Journal of Lipid Research) in a thesis or dissertation at no cost and with no explicit permission needed. Please see our copyright permissions page on the journal site for more information.

Best wishes,

Sarah Crespi

American Society for Biochemistry and Molecular Biology
11200 Rockville Pike, Rockville, MD
Suite 302
240-283-6616
JBC | MCP | JLR

From: **Frontiers in Genetics Editorial Office** genetics.editorial.office@frontiersin.org
Subject: Re: Reprints permission
Date: October 19, 2016 at 1:13 AM
To: rab501@mail.usask.ca

FI

Dear Mr Bi,

Thank you very much for your message.

As this is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY), the use, distribution or reproduction is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

In regards to the reuse of figures we always recommend that you seek the permission from the original authors so that no third-party copyright licenses apply.

Should you have any further questions please do not hesitate to get in touch.

Best wishes,

Sabrina

Sabrina Lehner, PhD
Review Operations Assistant

Frontiers | Editorial Office - Collaborative Peer Review Team
Review Operations Manager: Judyta Sorokowska

Frontiers
www.frontiersin.org
EPFL Innovation Square, Building I
Lausanne, Switzerland
Office T +41 21 510 1740

Frontiers community journals rapidly rise to become the most cited open-access journals in their fields.

For technical issues, please visit our Frontiers Help Center frontiers.zendesk.com

On Tue, Oct 18, 2016 at 7:44 AM, Bi, Ran (Frontiers Zendesk) <support@frontiersin.org> wrote:

Please type your reply above this line

frontiers Helpdesk support@frontiersin.org

You are registered as a CC on this support request (55074). Reply to this email to add a comment to the request.

Bi, Ran
Oct 18, 07:44 CEST

Thank you for your help!
Ran

???? iPhone

? 2016?10?17????11:42?Janos (Frontiers Zendesk) <support@frontiersin.org< mailto:support@frontiersin.org>> ???

Janos (Frontiers Zendesk)
Oct 18, 07:41 CEST

Ran Bi <rab501@mail.usask.ca>

Dear Mr Bi,

Thank you for your email. I am forwarding your request to my colleagues at the Editorial Office (in CC), who will get in touch with you soon.

Best regards,

Janos Eszter

Senior IT Helpdesk Analyst

Bi, Ran
Oct 17, 21:06 CEST

Dear Sir/Madam:

My name is Ran Bi, I am a student in University of Saskatchewan (Canada) working on my PhD thesis. I would like to ask for your permission to use “Figure 2. The biogenesis of microRNAs.” from Jung HJ and Suh Y. (2015). Regulation of IGF -1 signaling by microRNAs. Front Genet. 5:472. in my PhD thesis titled: EPIGENETIC REGULATION OF Hoxa1 AND Hoxa2.

Thank you!

Ran Bi
PhD Student
College of Pharmacy and Nutrition
University of Saskatchewan
107 Wiggins Road, Health Sciences Building, Room GD10
Saskatoon, Saskatchewan, S7N 5E5, Canada
Telephone (306) 966-6357

Please visit the Frontiers Help Center.

From: **Yousin Suh** yousin.suh@einstein.yu.edu
Subject: RE: Figure reprint request
Date: November 7, 2016 at 10:42 AM
To: Bi, Ran rab501@mail.usask.ca
Cc: Hwa Jin Jung hwa-jin.jung@einstein.yu.edu

YS

Hi Ran, it is OK as long as you properly indicate the source of the Figure, which I am sure you will. Yousin

---

From: Bi, Ran [mailto:rab501@mail.usask.ca]
Sent: Friday, November 4, 2016 9:19 AM
To: Yousin Suh <yousin.suh@einstein.yu.edu>
Subject: Figure reprint request

Dear Ms Suh:

My name is Ran Bi, I am a student in University of Saskatchewan (Canada) working on my PhD thesis. I am emailing you because I would like to ask for your permission to use “Figure 2. The biogenesis of microRNAs.” from your publication: Jung HJ and Suh Y. (2015). Regulation of IGF -1 signaling by microRNAs. Front Genet. 5:472. in my PhD thesis titled: EPIGENETIC REGULATION OF Hoxa1 AND Hoxa2. I found it a very nice figure to clearly illustrate the biogenesis of microRNAs and it would be a great support to my literature review of miRNA in my PhD thesis. I have already contacted the publisher and they suggested me to also ask you for permission to use the figure. So I sincerely hope you can grant me the permission to reprint this figure in my PhD thesis.

Thank you in advance for your help!

Sincerely,
Ran Bi
PhD Student
College of Pharmacy and Nutrition
University of Saskatchewan
107 Wiggins Road, Health Sciences Building, Room GD10
Saskatoon, Saskatchewan, S7N 5E5, Canada
Telephone (306) 966-6357

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Jan 31, 2017

This Agreement between Ran Bi ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4039690874029
License date	Jan 31, 2017
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Developmental Dynamics
Licensed Content Title	Hoxa2 plays a direct role in murine palate development
Licensed Content Author	Tara M. Smith,Xia Wang,Wei Zhang,William Kulyk,Adil J. Nazarali
Licensed Content Date	Aug 3, 2009
Licensed Content Pages	10
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Title of your thesis / dissertation	EPIGENETIC REGULATION OF Hoxa1 AND Hoxa2
Expected completion date	Jan 2017
Expected size (number of pages)	250
Requestor Location	Ran Bi College of Pharmacy and Nutrition University of Saskatchewan Saskatoon, SK S7N 5E5 Canada Attn: Ran Bi
Publisher Tax ID	EU826007151
Billing Type	Invoice
Billing Address	Ran Bi College of Pharmacy and Nutrition University of Saskatchewn Saskatoon, SK S7N 5E5 Canada